



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 157801**

**TO: Celine Qian**  
**Location: REM/2A64/2C70**  
**Art Unit: 1636**  
**Tuesday, July 05, 2005**

**Case Serial Number: 09/936271**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**THIS PAGE BLANK (USPTO)**





# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or *contact:*

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**THIS PAGE BLANK (USPTO)**

ATTN: EX MRB

CRFB

Access DB# 157801

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 6/28/05  
Art Unit: 1636 Phone Number 301-2-0777 Serial Number: 09/936271  
Mail Box and Bldg/Room Location: 2A 64 Results Format Preferred (circle): CARD DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of invention: Novel Human Kallikrein-like genes

Inventors (please provide full names): Yousef et al.

Earliest Priority Filing Date: 9/10/2004

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

search SEQ ID NO: 13 : 1 - 5,000 bp  
NA 11570 6000 - 7000 bp  
8000 - 9000 bp  
10000 - 11,570 bp

search SEQ ID NO: 14

AA 293  
both commercial database & interference.

ME

## STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>4</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) <u>1</u>	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>7/1/05</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>7/5/05</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. Review Time _____	Fulltext _____	Sequence Systems <u>DBA</u> <u>DBA</u>
Clerical Prep. time: _____	Patent Family _____	WWW/Internet _____
Online Time _____	Other _____	Other (specify) _____

FTO-1590 (8-01)

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 4527.92 Seconds

(without alignments)  
16811.954 Million cell updates/sec

Title: US-09-936-271c-13\_COPY\_10000\_11570

Perfect score: 1571

Sequence: 1 agggaggtatgggaattga.....cccgagaataaactgagaag 1571

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1571	100.0	11570	6	CQ788219	Sequence
2	1571	100.0	11570	6	CQ874885	Sequence
3	1571	100.0	11570	6	CQ874961	Sequence
4	1571	100.0	11570	9	AF135028	Homo sapi
5	1571	100.0	107487	9	AC011483	Homo sapi
6	1571	100.0	230000	9	AF243527	Homo sapi
7	1555.2	99.0	217346	2	AC027602	Homo sapi
8	1482	94.3	200792	2	AC130782	Pan trogl
9	482	30.7	586	6	AX331407	Sequence
10	482	30.7	586	6	AX331824	Sequence
11	482	30.7	586	11	G41906	SHGC-56840
12	482	30.7	735	6	AX429955	Sequence
13	482	30.7	738	6	BD139877	Compounds
14	482	30.7	1143	6	CQ720600	Sequence
15	482	30.7	1301	9	AY461805	Homo sapi
16	482	30.7	1370	9	AY279380	Homo sapi
17	482	30.7	1381	6	BD107879	Homo sapi
18	482	30.7	1387	6	CQ874884	Sequence
19	482	30.7	1387	6	CQ874960	Sequence

20	482	30.7	1387	9	AF168768	Homo sapi
21	482	30.7	1438	9	AY279381	Homo sapi
22	482	30.7	1439	6	BD107880	36 human
23	482	30.7	1499	6	AR352504	Sequence
24	482	30.7	1499	6	BD005362	Protease.
25	482	30.7	1516	6	BD107865	36 human
26	482	30.7	1527	9	BC008036	Homo sapi
27	482	30.7	1570	6	AR252580	Sequence
28	482	30.7	1570	6	AR528679	Sequence
29	482	30.7	1570	6	AX080829	Sequence
30	482	30.7	1570	6	AX403421	Sequence
31	482	30.7	1570	6	AX464322	Sequence
32	482	30.7	1570	9	AY359010	Homo sapi
33	476	30.3	1476	6	AR078184	Sequence
34	476	30.3	1476	6	AR137506	Sequence
35	476	30.3	1476	6	AR242358	Sequence
36	476	30.3	1476	6	BD082659	Sequence
37	460	29.3	1504	6	AR263823	Sequence
38	449	28.6	673	6	AX067353	Sequence
39	315.4	20.1	498	11	BV209338	KLK5_3564
40	236.6	15.1	321519	2	AL714004	Homo sapi
41	236	15.0	175559	9	AC112498	Homo sapi
42	234.4	14.9	174380	2	AC016075	Homo sapi
43	233.2	14.8	151880	9	AP005716	Homo sapi
44	231.2	14.7	79516	9	AC004834	Homo sapi
45	231.2	14.7	156949	9	AL161652	Human DNA

#### ALIGNMENTS

RESULT 1

CQ788219

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .11570

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

ORIGIN

Query Match	100.0%;	Score	1571;	DB	6;	Length	11570;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1571;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	AGGGAGGTATGGGAATTGAAGACAGGAAACACAAATTAGTCCAAAGCGAATGGATTCTA	60				
Db	10000	AGGGAGGTATGGGAATTGAAGACAGGAAACACAAATTAGTCCAAAGCGAATGGATTCTA	10059				
Qy	61	TTGGGAGTGAATTCGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	120				
Db	10060	TTGGGAGTGAATTCGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	10119				
Qy	121	CARCTATATGAGGGGCAATTAATGGATAGATGCAAGTGTCTGTTCAAC	180				
Db	10120	CAACTATATGAGGGGCAATTAATGGATAGATGCAAGTGTCTGTTCAAC	10179				
Qy	181	ATGCTATGATGCACAGGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	240				
Db	10180	ATGCTATGATGCACAGGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC	10239				





Qy	541	AGTGA	CAAAATGGGGTCTTAAAGTTGAACTTGGAGGCCAGGCATGTGGCTCAGCGCTGTA	600			
Db	10540	AGTGA	CAAAATGGGGTCTTAAAGTTGAACTTGGAGGCCAGGCATGTGGCTCAGCGCTGTA	10599			
Qy	601	ATCCCA	ACAATTTGGAGGCTGAGGTGGGCGAATCACTTGAGGCCAGGAGTTTCGAGACCCAG	660			
Db	10600	ATCCCA	ACAATTTGGAGGCTGAGGTGGGCGAATCACTTGAGGCCAGGAGTTTCGAGACCCAG	10659			
Qy	661	CTTGGC	CAACATGGTGAACCACCCGCTCTCTACAAAAAATACAAAAAATTAGCCGGGTGT	720			
Db	10660	CTTGGC	CAACATGGTGAACCACCCGCTCTCTACAAAAAATACAAAAAATTAGCCGGGTGT	10719			
Qy	721	GGTGAT	GGACAACCTGTAGTCA	CAGCTPACTTGGGAGCTGAGCAGGAGAAATTGCTTGAAC	780		
Db	10720	GGTGAT	GGACAACCTGTAGTCA	CAGCTPACTTGGGAGCTGAGCAGGAGAAATTGCTTGAAC	10779		
Qy	781	CCGGGAGAT	GGAGGCTG	CAGTGAAGTCA	GCGGCTC	CAACCTGGGCAACA	840
Db	10780	CCGGGAGAT	GGAGGCTG	CAGTGAAGTCA	GCGGCTC	CAACCTGGGCAACA	10839
Qy	841	GAGTAAG	ACTCCATCTCAAAAAA	AAAAAGCTGGATTTGGAGTGAATAATTAATAACATT	900		
Db	10840	GAGTAAG	ACTCCATCTCAAAAAA	AAAAAGCTGGATTTGGAGTGAATAATTAATAACATT	10899		
Qy	901	CTCCCTCT	CTCTCTTTTGGCTGTGCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT	960			
Db	10900	CTCCCTCT	CTCTCTTTTGGCTGTGCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT	10959			
Qy	961	CTGTACTTTT	CCATCTCTGTGTCTGTTCCCATCTGTCTTCCATCTATGGGCATCTCTG	1020			
Db	10960	CTGTACTTTT	CCATCTCTGTGTCTGTTCCCATCTGTCTTCCATCTATGGGCATCTCTG	11019			
Qy	1021	GGTCTCT	CAATCTCTCTTGCCACTTTGGCCACATCTCTGCTCTCTCATGSCCCCCCTT	1080			
Db	11020	GGTCTCT	CAATCTCTCTTGCCACTTTGGCCACATCTCTGCTCTCTCATGSCCCCCCTT	11079			
Qy	1081	TCCTCTCT	GCAGGGTGAATTC	TGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGAATC	1140		
Db	11080	TCCTCTCT	GCAGGGTGAATTC	TGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGAATC	11139		
Qy	1141	GTGTCTCT	GGGAGATTAACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGAACTCTC	1200			
Db	11140	GTGTCTCT	GGGAGATTAACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGAACTCTC	11199			
Qy	1201	TGCAAGTT	CACAAAGTGGATCCAGGAAACATCCAGGCCCAACTCTGAGTCATCCCGGA	1260			
Db	11200	TGCAAGTT	CACAAAGTGGATCCAGGAAACATCCAGGCCCAACTCTGAGTCATCCCGGA	11259			
Qy	1261	CTCAGCA	CCGGCATCCCACTGTCTGCAGGAGACGCCCTGACACTCTCTTTCAGACCT	1320			
Db	11260	CTCAGCA	CCGGCATCCCACTGTCTGCAGGAGACGCCCTGACACTCTCTTTCAGACCT	11319			
Qy	1321	CATTCTCT	CCAGAGATTTAGAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCT	1380			
Db	11320	CATTCTCT	CCAGAGATTTAGAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCT	11379			
Qy	1381	GGACTCA	GGGTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTGGG	1440			
Db	11380	GGACTCA	GGGTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTTGAACCTGGG	11439			
Qy	1441	AACAATTT	CAAACTGTCACAGGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCA	1500			
Db	11440	AACAATTT	CAAACTGTCACAGGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCA	11499			
Qy	1501	TCCTCAAG	CTCAGGCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAT	1560			
Db	11500	TCCTCAAG	CTCAGGCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAT	11559			
Qy	1561	AAACTG	AGAAG	1571			
Db	11560	AAACTG	AGAAG	11570			

RESULT 4	AF135028	11570 bp	DNA	linear	PRI 27-JUN-2000
LOCUS	AF135028	11570 bp	DNA	linear	PRI 27-JUN-2000
DEFINITION	AF135028	11570 bp	DNA	linear	PRI 27-JUN-2000
ACCESSION	AF135028	11570 bp	DNA	linear	PRI 27-JUN-2000
VERSION	AF135028.1	GI:4589282			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					



Db	10180	ATGCTATGATGCAACGGCAGGCCTCCAAACAACCATTATTCACAGTTCAGATGCCCAC	10239
Qy	241	AGTGCCCGAGATCAGAGAACCTCATCCAGGGCTGAGAACCCTGATTTTGTTCAGAAGGGAG	300
Db	10240	AGTGCCCGAGATCGAGGAACCTCATCCAGGGCTGAGAACCGTATTTTGTTCAGAAGGGAG	10299
Qy	301	GTAATAAGATGGGTTCGTGGAGATGGGGAAGAAAGGTGTGTCTCAGTAAGAGAAAATAA	360
Db	10300	GTAATAAGATGGGTTCGTGGAGATGGGGAAGAAAGGTGTGTCTCAGTAAGAGAAAATAA	10359
Qy	361	GGCCTCGACAGGCTGAGGGGAGGTGACAGAAAGGAGCGCAGAGATACACGATGA	420
Db	10360	GGCCTCGACAGGCTGAGGGGAGGTGACAGAAAGGAGCGCAGAGATACACGATGA	10419
Qy	421	GGGAGACAGCTCGAAACAGAAAGTAGACGAAAGATTCCAGATGTGGAGAGGAAGGGTCA	480
Db	10420	GGGAGACAGCTCGAAACAGAAAGTAGACGAAAGATTCCAGATGTGGAGAGGAAGGGTCA	10479
Qy	481	CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAAGAGAAATGGAGTGGAG	540
Db	10480	CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAAGAGAAATGGAGTGGAG	10539
Qy	541	AGTGACAAATGGGGTCTAAAGGTTGAACTTTGAGGCGCAGGCATGTGGCTCAGCGCTGTA	600
Db	10540	AGTGACAAATGGGGTCTAAAGGTTGAACTTTGAGGCGCAGGCATGTGGCTCAGCGCTGTA	10599
Qy	601	ATCCCACAACCTTTGGAGGCTGAGGTGGGCGAATCACTTGAGGCCAGGAGTTTCGAGACCAG	660
Db	10600	ATCCCACAACCTTTGGAGGCTGAGGTGGGCGAATCACTTGAGGCCAGGAGTTTCGAGACCAG	10659
Qy	661	CTGGGCCAAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAAATTAGCCGGGTGT	720
Db	10660	CTGGGCCAAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAAATTAGCCGGGTGT	10719
Qy	721	GGTGATGGACACTGTAGTCAACAGCTACTTTGGAGGCTGAGCAGAGAAATTTGTTGAAC	780
Db	10720	GGTGATGGACACTGTAGTCAACAGCTACTTTGGAGGCTGAGCAGAGAAATTTGTTGAAC	10779
Qy	781	CCGGGAGATGGAGGCTGACGTGAGCTCAGGCGCACTGGCGTCCAACCTTGGGCAACA	840
Db	10780	CCGGGAGATGGAGGCTGACGTGAGCTCAGGCGCACTGGCGTCCAACCTTGGGCAACA	10839
Qy	841	GAGTAAGACTCCATCTCAAAAAAAAAGCTGGATTTGGAGTGAATATTAATAACATT	900
Db	10840	GAGTAAGACTCCATCTCAAAAAAAAAGCTGGATTTGGAGTGAATATTAATAACATT	10899
Qy	901	CTCCCTCTCTCCTTTTGGCTGTCTCACTCTGTCTTTTTCTGCAATTTCTTCATCT	960
Db	10900	CTCCCTCTCTCCTTTTGGCTGTCTCACTCTGTCTTTTTCTGCAATTTCTTCATCT	10959
Qy	961	CTGTACTTTTCCATCTCTGTGTGTCTTCCCATCTGCTTCCATCTATGGCATCTCTG	1020
Db	10960	CTGTACTTTTCCATCTCTGTGTGTCTTCCCATCTGCTTCCATCTATGGCATCTCTG	11019
Qy	1021	GGTCTCTCATGTCTCCTTTCTGCCCACTTTGGCCACATCTCTGCTCTCTCATGCCCCCTT	1080
Db	11020	GGTCTCTCATGTCTCCTTTCTGCCCACTTTGGCCACATCTCTGCTCTCTCATGCCCCCTT	11079
Qy	1081	TCTCTCTCGAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGGTCCTCGAGGACTC	1140
Db	11080	TCTCTCTCGAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGGTCCTCGAGGACTC	11139
Qy	1141	GTGTCTCGGGAGATTTACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGAACCTC	1200
Db	11140	GTGTCTCGGGAGATTTACCTTTGTGCCGGCCCAACAGACCGGGTGTCTACAGAACCTC	11199
Qy	1201	TGCAAGTTTACCAGTGGATCCAGGAAAACATCCAGGCCAACTCTGTAGTCTATCCAGGA	1260
Db	11200	TGCAAGTTTACCAGTGGATCCAGGAAAACATCCAGGCCAACTCTGTAGTCTATCCAGGA	11259
Qy	1261	CTCAGCACACCGGCATCCCACTCTGTGAGGAGACGCCCTGACACTCTCTTTTTCAGACCT	1320

Db	11260	CTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCTCTTTTCAGACCCCT	11319
Qy	1321	CATTCTCTCCAGAGATGTTGAGAATGTTTCATCTCTCCAGGCCCTGACCCCAATGTCCTCTCT	1380
Db	11320	CATTCTCTCCAGAGATGTTGAGAATGTTTCATCTCTCCAGGCCCTGACCCCAATGTCCTCTCT	11379
Qy	1381	GGACTCAGGCTCTGCTTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG	1440
Db	11380	GGACTCAGGCTCTGCTTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG	11439
Qy	1441	AACAAATTTCAAAATCTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCCCTGGGCACTTTTCA	1500
Db	11440	AACAAATTTCAAAATCTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCCCTGGGCACTTTTCA	114999
Qy	1501	TCCTCAAGCTCTCAGGGCCCATCTCCCTCTCTGTCAGCTCTGACCCCAAATTTTAGTCCCAAGAAT	1560
Db	11500	TCCTCAAGCTCTCAGGGCCCATCTCCCTCTCTGTCAGCTCTGACCCCAAATTTTAGTCCCAAGAAT	11559
Qy	1561	AAACTGAGAAG	1571
Db	11560	AAACTGAGAAG	11570

```

RESULT 5
AC011483/c
LOCUS
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
ACCESSION AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.

```

```

Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.

FEATURES
    source
        1..107487
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /clone="CTB-147C22"
            /998..65494
            /note="NOTE: Shatter libraries failed to resolve
            dinucleotide repeat. Unsure number of repeat copies
            64998-65494. Forced join 65015."

    misc_feature

```

Query Match	100.0%;	Score 1571;	DB 9;	Length 107487;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1571:	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;



CDS  
/gene="KLK1"  
/product="renal kallikrein"  
complement(join(6963..7118,7668..7804,7923..8212,  
9482..9641,11472..11517))  
/gene="KLK1"  
/notes="serine protease"  
/codon\_start=1  
/product="renal kallikrein"  
/protein\_id="AAG33353.1"  
/db\_xref="GI:11244758"  
mRNA  
/translation="MMFLVLCALSLSGTGGAAPPIQSRIVGVGWECEKHSQPWQVLAALYH  
FSTFQCGIIVHQWLTAAHCISDNYQLWGRHNLFDDENTAQFVHVSFPHPGFN  
MSLLENHROADSDYSHDLMLLTPADTITDAVKVELPTPEPEVSGTCLASGWS  
IEPENSFPDLQCDLKIILNDECKKAHVQKVTDEMLCVHLEGGKDTCTGDSGGPL  
MCGVLOGVTSWGVPCGTENKPSVAVRVLSYKWIHDTIAENS"  
complement(join(<13552..13704,14377..14530,14678..14917,  
15416..15569,19204..19246))  
/product="ACO protease"  
complement(join(13552..13704,14377..14530,14678..14917,  
15416..15569,19204..19246))  
/notes="serine protease"  
/codon\_start=1  
/product="ACO protease"  
/protein\_id="AAG33354.1"  
/db\_xref="GI:11244759"  
/translation="MWLLLTSLFLLASTAODGDKLLEGDECAPHSQPWQVVALYERGR  
FNGASLISPHWLSAAHCSRFMRVRLGHNLRKRDGPQLRTTSRVIPHPRYEARS  
HRNDIMLLRLVQPARLNPQVPAVLPTPCPHGECACVSGWGLVPLSSVSPDPTLHC  
ANISLISDTSKDSYFRLTNTWVCAGAEAGSCEBGDSGGPLVCGGILQGVISWGD  
VPCDNTTKPGVYTKVCHYLEWIRETKRN"  
42595..>47769  
/gene="KLK3"  
join(<42595..42640,43880..44039,45669..45955,46099..46235,  
47614..>47769)  
/gene="KLK3"  
/product="prostate specific antigen"  
join(42595..42640,43880..44039,45669..45955,46099..46235,  
47614..>47769)  
/gene="KLK3"  
/notes="serine protease"  
/codon\_start=1  
/product="prostate specific antigen"  
/protein\_id="AAG33355.1"  
/db\_xref="GI:11244760"  
/translation="MMVVFVFLTSVTWIGAAPLISRIVGVGWECEKHSQPWQVLAAS  
RGRVACGVVLVHPQWLTAAHCIRNKSIVILLGRHSLFHPEDTQGVQVSHSFPHPLYD  
MSLLKNRFLRPDSDSHDLMLRLSEPAELTDAVKVNDLPTQBPALGTTTCYASGWGSI  
EPFEFLTPKLQCVDLHVISNDVCAOVHPQKVKFMLCAGRWTGKSTCGSDSGGPLV  
CNGVLOGITSWSEPCALPERPSLYTKVHYRWIKDTIVANP"  
61139..>66229  
/gene="KLK2"  
join(<61139..61184,62391..62550,64142..64428,64542..64678,  
66074..>66229)  
/gene="KLK2"  
/product="glandular kallikrein"  
join(61139..61184,62391..62550,64142..64428,64542..64678,  
66074..>66229)  
/gene="KLK2"  
/notes="serine protease"  
/codon\_start=1  
/product="glandular kallikrein"  
/protein\_id="AAG33356.1"  
/db\_xref="GI:11244761"  
/translation="MMWDLVLSIALSVGTGAVPLIQSRIVGVGWECEKHSQPWQVAVVS  
HGWAHCGVLVHPQWLTAAHCLKNSQVWLGRHNLPEPEDTGORVPVSHSFPHLYN  
MSLLKHOSLPDENSDSHDLMLRLSEPAKITDVVKVGLPTQBPALGTTTCYASGWGSI  
EPFEFLRPSLQCVSLHLLNDMCARAYSEKVFTEMLCAGLWGTGGKDTCCGDSGGPLV  
CNGVLOGITSWSEPCALPEKPAVYTKVHYRWIKDTIAANP"  
complement(join(<94602..94754,96027..96163,96247..96497,  
96919..97081,98345..>98405))  
/product="serine protease"  
complement(join(94602..94754,96027..96163,96247..96497,  
96919..97081,98345..>98405))

gene  
/codon\_start=1  
/product="serine protease"  
/protein\_id="AAG33357.1"  
/db\_xref="GI:11244762"  
/translation="MATAGNPMGWFLGYLILVAGSLVSGSCSIINGEDCSPHSQPW  
QAALVMENELFCGSLVLPQWVLSAAHCFQNSYITGLHLSLEADQEPGSMYEAALS  
VRHPEYRPLLANDMLIILDESVSSEDTSIRISIASQCFTAGNSCLVSGWGLLANGR  
MPTVLOCVNVSVSEVSCSLYDPLYPHSMFCAGGHDQKDCSNGSGGGLICNGYLO  
GLVSGKAPCGQGVGVYTNLCKFTETWIKVQAS"  
complement(join(131301..>140303))  
/gene="KLK5"  
/notes="synonym: SCTE"  
complement(join(<131301..131456,136310..136443,  
136529..136785,137525..137690,140255..>140303))  
/gene="KLK5"  
/product="stratum corneum trypsin-like serine protease"  
complement(join(131301..131456,136310..136443,  
136529..136785,137525..137690,140255..>140303))  
/gene="KLK5"  
/codon\_start=1  
/product="stratum corneum trypsin-like serine protease"  
/protein\_id="AAG33358.1"  
/db\_xref="GI:11244763"  
/translation="MMVLCALITALLGYTDARSDDSSRIINGSDCDMMHTQPWQAAL  
LIRPNOLYCGAVLHPQWLTAAHCKKVRVRLGHVLSLPVYESGQMFQGVKSIPIH  
PGYSHCHSDNMLIILKNRIRPTKDVIRNVSSHCPBSAGTKCLVSGWGTTKSPQVHF  
PKVQLCINLSVLSQKRCEDAYPQIDDTMFCAGDKAGRDCQGDGSGVPPVVCNGSLQGL  
VSMGDPYPCARPNNRPVYTNLCKFTWIKVQIQAENS"  
complement(join(<146834..146986,149628..149764,  
151186..151433,155052..155208,155948..>155987))  
/product="protease M"  
complement(join(146834..146986,149628..149764,  
151186..151433,155052..155208,155948..>155987))  
/notes="serine protease; also called neurosin or zyme"  
/codon\_start=1  
/product="protease M"  
/protein\_id="AAG33359.1"  
/db\_xref="GI:11244764"  
/translation="MKKLMVLSLIAAAWAEQNKLVHGGPCDKTSHPHYQAALYTSQH  
LLCGVLIHPLWLTAHCKKNLQVFLGKNLRQRESQESQSVVRVAVIHPDYDAAS  
HDODIMLLRLARPAKLSLTIPLPLERDCSANTTSCHILGWKGTDFPDFTIQCAVI  
HLVSRCECHAVPGQITONMLCAGDEKYGKDCQGDGSGGPLVCGDHLRLGLVSGNIPFC  
GSKEKGVYTNVCRVYTNLCKFTWIKVQIQAENS"  
complement(join(<165420..>170283))  
/gene="KLK7"  
/notes="synonym: SCCE"  
complement(join(<165420..165575,167672..167808,  
168124..168371,169651..169798,170211..>170283))  
/gene="KLK7"  
/product="stratum corneum chymotryptic enzyme"  
complement(join(165420..165575,167672..167808,  
168124..168371,169651..169798,170211..>170283))  
/gene="KLK7"  
/notes="serine protease"  
/codon\_start=1  
/product="stratum corneum chymotryptic enzyme"  
/protein\_id="AAG33360.1"  
/db\_xref="GI:11244765"  
/translation="MARSLLLPLQILLSLALETAGEAOGDKIIDGAPCARGSHPMQ  
VALLSGNLHCGGLVNERWLTAAHCKNEYTVHLGSDTLGDRRAQRIKASKSRHP  
GYSTQTHVNDLMLVKNLSQARLSNNVKKVRLPSRCPPTCTVSGMGTTSPDVTFP  
VSMGTFCQGNLIPDQCTKVYKDLLENMLCAGIDPDRSKKNACNDSGSGPLVCRGLTQGL  
SDMLGTFCQGNLIPDQCTKVYKDLLENMLCAGIDPDRSKKNACNDSGSGPLVCRGLTQGL  
complement(join(<183943..184098,185635..185768,  
187865..188127,188293..188452,188967..>189036))  
/product="neuropsin"  
complement(join(183943..184098,185635..185768,  
187865..188127,188293..188452,188967..>189036))  
/notes="serine protease; also called ovasin"  
/codon\_start=1  
/product="neuropsin"  
/protein\_id="AAG33361.1"  
/db\_xref="GI:11244766"

		/translation="MGRPRPRAAKTWMLFLLGGAWAGHSRAQEDKVLGGHECPHQSP PWAALFQGGQLLGGVLYGVGNWLVTAACHCKPKYKTVRLGDHSLQNKDGPQIPVQV SIHPHCVNSDSDVNDHDLMLQLRDOASLGSKVPISLADHCTQPKQCTVSGWGT TSRENPFDTLNCAEVKIFPKKCEDAYPGQITDGMVACAGSSKGAUTCCQDSGGPLVC DGAQGITWSGSPCGRSDRPGVYTNICRYLDWIKLIIGSKG" complement(join(<i>190980..191129,191573..191709, 194324..194589,197048..197204,197370..>197412)) /product="kallikrein-like 3" /note="kallikrein-like 3" 194324..194589,197048..197204,197370..197412)) /note="serine protease" /codon_start=1 /product="kallikrein-like 3" /protein_id="AAG33362.1" /db_xref="GI:11244767" /translation="MKLIGLCALLSLLAGHWADTRAIGAEFCRPNSPHQAGLFHLT RLFCGATLISDRWLLTAHCRKPYLVRIGEHHLWKVEGEQFLRVDTPFPHGFNPD LSANDHDDIMLRPLQARLSPAVQPLNUSQTCVSPQMCLISGHWAVSSPKALFPV TLQCANISILENKLHWAYPGHISDSMLCAGLWEGGRSCQSDSGGPLVCGNGLAGVV SGNAEPCSRPRRPAVTVSVCHYLDWIQEIEMEN"			
Query Match		100.0%;	Score 1571;	DB 9;	Length 230000;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1571;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAAATTAGTCCAAAGCAATGGATTCTTA	60		
Db	132548	AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAAATTAGTCCAAAGCAATGGATTCTTA	132489		
Qy	61	TTGGGAGTGATTCGCCCCTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA	120		
Db	132488	TTGGGAGTGATTCGCCCCTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA	132429		
Qy	121	CAACTATATGGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAGTGTGCTGTTCAAC	180		
Db	132428	CAACTATATGGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAGTGTGCTGTTCAAC	132369		
Qy	181	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATATCCAGCTTCAGATGCCAC	240		
Db	132368	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATATCCAGCTTCAGATGCCAC	132309		
Qy	241	AGTGCCAGATCGAGGAACCTCATCCAGGGCTGAGAACCGTATTTTCCAGAAGGGAG	300		
Db	132308	AGTGCCAGATCGAGGAACCTCATCCAGGGCTGAGAACCGTATTTTCCAGAAGGGAG	132249		
Qy	301	GTATAAGGATGGGTGGTGAGATGGGAGGAGGTGTGTCCTCAAGTAAGGAATAA	360		
Db	132248	GTATAAGGATGGGTGGTGAGATGGGAGGAGGTGTGTCCTCAAGTAAGGAATAA	132189		
Qy	361	GGCTGCACAGGCTGGAGGGGAGAGTGCAGAGAGAAAGGAGGCGGAGAGATACACGATGA	420		
Db	132188	GGCTGCACAGGCTGGAGGGGAGAGTGCAGAGAGAAAGGAGGCGGAGAGATACACGATGA	132129		
Qy	421	GGGAGACAGGCTGAAACAGAAAGTAGAGACAGAGATTCGAGATGTGGAGAGGAGGTCA	480		
Db	132128	GGGAGACAGGCTGAAACAGAAAGTAGAGACAGAGATTCGAGATGTGGAGAGGAGGTCA	132069		
Qy	481	CAGACCCCGGAAATCATGTGTGGACAAACAGGAATCTGGAGAGGAGAGATGGAGTGGAG	540		
Db	132068	CAGACCCCGGAAATCATGTGTGGACAAACAGGAATCTGGAGAGGAGAGATGGAGTGGAG	132009		
Qy	541	AGTGACAAATGGGCTTAAAGGTTGAACCTTGAGGCGCAGGCATGGTGGCTCACGCCCTGTA	600		
Db	132008	AGTGACAAATGGGCTTAAAGGTTGAACCTTGAGGCGCAGGCATGGTGGCTCACGCCCTGTA	131949		
Qy	601	ATCCCAACATTTGGAGGCTGAGTGGCGGAATCACTTGAGGCGCAGAGATTTCGAGACCAG	660		
Db	131948	ATCCCAACATTTGGAGGCTGAGTGGCGGAATCACTTGAGGCGCAGAGATTTCGAGACCAG	131889		
Qy	661	CCTGGCCCAACATGGTGAACCCCGTCTACAAAAAAATACAAAAAATTAGCCGGGTGT	720		
Db	131888	CCTGGCCCAACATGGTGAACCCCGTCTCTACAAAAAAATACAAAAAATTAGCCGGGTGT	131829		

Qy	721	GGTGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTCAGGAGAGAAATGCTTGAAC	780
Db	131828	GGTGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTCAGGAGAGAAATGCTTGAAC	131769
Qy	781	COGGAGATGGAGGCTGCAGTGAGCTGAGGTCAAGGCCATCTGCCTCCAACTGGGCAACA	840
Db	131768	COGGAGATGGAGGCTGCAGTGAGCTGAGGTCAAGGCCATCTGCCTCCAACTGGGCAACA	131709
Qy	841	GAGTAAGACTCCCATCTCAAAAAAAGCTGAGATTGGAGTGAAATATTAATAACATT	900
Db	131708	GAGTAAGACTCCCATCTCAAAAAAAGCTGAGATTGGAGTGAAATATTAATAACATT	131649
Qy	901	CTCCCT	960
Db	131648	CTCCCT	131589
Qy	961	CTGTACTTTCATCTCTGTGTGTCTGTGTGTCTGTGTGTCTGTGTGTCTGTGTGTCTGTG	1020
Db	131588	CTGTACTTTCATCTCTGTGTGTCTGTGTGTCTGTGTGTCTGTGTGTCTGTGTGTCTGTG	131529
Qy	1021	GGTCTCTCATGTCT	1080
Db	131528	GGTCTCTCATGTCT	131469
Qy	1081	TCTCTCTCTCAGGCTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTC	1140
Db	131468	TCTCTCTCTCAGGCTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTC	131409
Qy	1141	GTGTCTCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTC	1200
Db	131408	GTGTCTCTGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTC	131349
Qy	1201	TGCAAGTTTCCAAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGA	1260
Db	131348	TGCAAGTTTCCAAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGA	131289
Qy	1261	CTCAGCACACCGGCATCCCCACCTGTCTGCAGGAGACGCCCTGACACTCTCTTTCAGACCT	1320
Db	131288	CTCAGCACACCGGCATCCCCACCTGTCTGCAGGAGACGCCCTGACACTCTCTTTCAGACCT	131229
Qy	1321	CATTCTCTCCAGAGATGTTGAGAAATGTTCAATCTCTCAGGCCCTGACCCCATGTCTCT	1380
Db	131228	CATTCTCTCCAGAGATGTTGAGAAATGTTCAATCTCTCAGGCCCTGACCCCATGTCTCTCT	131169
Qy	1381	GGACTCAGGCTCTGCTTCCCCACATTTGGCTGACCGTGTCTCTCTAGTTTGAACCTGGG	1440
Db	131168	GGACTCAGGCTCTGCTTCCCCACATTTGGCTGACCGTGTCTCTCTAGTTTGAACCTGGG	131109
Qy	1441	AACAATTTCCAAACTGTCTCAGGGCGGGGTTCGGTCTCTCAATCTCCCTGGGGCACTTTCA	1500
Db	131108	AACAATTTCCAAACTGTCTCAGGGCGGGGTTCGGTCTCTCAATCTCCCTGGGGCACTTTCA	131049
Qy	1501	TCCTCAAGCTCAGGGCCCATTCCTTCTCTGACGCTTGACCCCAAAATTTAGTCCAGAAAT	1560
Db	131048	TCCTCAAGCTCAGGGCCCATTCCTTCTCTGACGCTTGACCCCAAAATTTAGTCCAGAAAT	130989
Qy	1561	AAACTCAGAAG 1571	
Db	130988	AAACTCAGAAG 130978	
RESULT 7			
AC027602/c		217346 bp	DNA linear
LOCUS		Homo sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT	
DEFINITION		SEQUENCE, 12 unordered pieces.	
AC027602		AC027602.4 GI:11178143	
ACCESSION		HTG; HTGS PHASE1; HTGS_DRAFT.	
VERSION		HTG; HTGS PHASE1; HTGS_DRAFT.	
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

## REFERENCE

1 (bases 1 to 217346)  
Birren,B., Linton,L., Nuebaum,C. and Lander,E.

## AUTHORS

Homosapiens chromosome 19, clone RP11-795B6

## JOURNAL

Unpublished

## REFERENCE

2 (bases 1 to 217346)

## AUTHORS

Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Cooke,A., Choepel,Y., Colangelo,W., Collins,S., Collymore,A., Castle,K., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

## TITLE

Direct Submission

## JOURNAL

Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

3 (bases 1 to 217346)

## AUTHORS

Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Bada,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,W., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severi,P., Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

## TITLE

Direct Submission

## JOURNAL

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On Nov 15, 2000 this sequence version replaced gi:11136831.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L9166

Center clone name: 795 B 6

----- Summary Statistics

Sequencing vector: M13; M7815; 31% of reads

Sequencing method: Plasmid; n/a; 69% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 210748 bases at least Q40

Consensus quality: 213655 bases at least Q30

Consensus quality: 215058 bases at least Q20

ORIGIN

Insert size: 194000; agarose-fp  
Insert size: 216246; sum-of-contigs  
Quality coverage: 11.9 in Q20 bases; agarose-fp  
Quality coverage: 10.7 in Q20.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 8149: contig of 8149 bp in length  
\* 8150 8249: gap of 100 bp  
\* 8250 9592: contig of 1343 bp in length  
\* 9593 9692: gap of 100 bp  
\* 9693 10733: contig of 1041 bp in length  
\* 10734 10833: gap of 100 bp  
\* 10834 13519: contig of 2686 bp in length  
\* 13520 13619: gap of 100 bp  
\* 13620 17510: contig of 3891 bp in length  
\* 17511 17610: gap of 100 bp  
\* 17611 24602: contig of 6992 bp in length  
\* 24603 24702: gap of 100 bp  
\* 24703 35434: contig of 10732 bp in length  
\* 35435 35534: gap of 100 bp  
\* 35535 124474: contig of 88940 bp in length  
\* 124475 124574: gap of 100 bp  
\* 124575 134664: contig of 10090 bp in length  
\* 134665 134764: gap of 100 bp  
\* 134765 162343: contig of 27579 bp in length  
\* 162344 162443: gap of 100 bp  
\* 162444 208917: contig of 46474 bp in length  
\* 208918 209017: gap of 100 bp  
\* 209018 217346: contig of 8329 bp in length.

## FEATURES

## source

1. 217346  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19"  
/clone="RP11-795B6"  
/clone\_lib="RPCI-11 Human Male BAC"

## misc\_feature

1. 8149  
/notes="assembly\_fragment"  
clone end:SP6  
vector\_side:left"

## misc\_feature

8250..9592  
/notes="assembly\_fragment"

## misc\_feature

9693..10733  
/notes="assembly\_fragment"

## misc\_feature

10834..13519  
/notes="assembly\_fragment"

## misc\_feature

13620..17510  
/notes="assembly\_fragment"

## misc\_feature

17611..24602  
/notes="assembly\_fragment"

## misc\_feature

24703..35434  
/notes="assembly\_fragment"

## misc\_feature

35535..124474  
/notes="assembly\_fragment"

## misc\_feature

124575..134664  
/notes="assembly\_fragment"

## misc\_feature

134765..162343  
/notes="assembly\_fragment"

## misc\_feature

162444..208917  
/notes="assembly\_fragment"

## misc\_feature

209018..217346  
/notes="assembly\_fragment"

## misc\_feature

clone end:T7  
vector\_side:right"

Query Match		99.0%; Score 1555.2; DB 2; Length 217346;
Best Local Similarity		99.7%; Pred. No. 0;
Matches 1568; Conservative		0; Mismatches 3; Indels 1; Gaps 1;
Qy	1	AGGAGGTATGGGAAATTGAAGACAGAGAAACACAAATTAAGTCCAAAGCAATGATTTCTA 60
Db	194423	AGGAGGTATGGGAAATTGAAGACAGAGAAACACAAATTAAGTCCAAAGCAATGATTTCTA 194364
Qy	61	TTGGAGTGATTTGCCCCCTAGAGACACTGGCNAATACCAAGGACACATTTTGGTTGTCA 120
Db	194363	TTGGAGTGATTTGCCCCCTAGAGACACTGGCNAATACCAAGGACACATTTTGGTTGTCA 194304
Qy	121	CAACTATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGTGCTTTCAAC 180
Db	194303	CAACTATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGTGCTTTCAAC 194244
Qy	181	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTTATCCAGCTTCAGATGCCAC 240
Db	194243	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTTATCCAGCTTCAGATGCCAC 194184
Qy	241	AGTCCCAGATCGAGGAACCTCATCCAGGGCTGAGAACGATATTTTTCAGAGGGAG 300
Db	194183	AGTCCCAGATCGAGGAACCTCATCCAGGGCTGAGAACGATATTTTTCAGAGGGAG 194124
Qy	301	GTATAAGGATGGGTTCGTGAGAAATGGGAAGGAGTGTGTCCAGTAAAGAAATAA 360
Db	194123	GTATAAGGATGGGTTCGTGAGAAATGGGAAGGAGTGTGTCCAGTAAAGAAATAA 194064
Qy	361	GGCTGTCACAGGCTGAGGGGAGAGTGCAGAGAAAGGAGGGGAGAGATACACGATGA 420
Db	194063	GGCTGTCACAGGCTGAGGGGAGAGTGCAGAGAAAGGAGGGGAGAGATACACGATGA 194004
Qy	421	GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTCGAGATGTGAGAGGAAGGTCA 480
Db	194003	GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTCGAGATGTGAGAGGAAGGTCA 193944
Qy	481	CAGACCCCCCGGAATGATGTGTGCAACACAGGAATCTGGAAGAGGAAGATGAGTGGAG 540
Db	193943	CAGACCCCCCGGAATGATGTGTGCAACACAGGAATCTGGAAGAGGAAGATGAGTGGAG 193884
Qy	541	AGTGACAAATGGGTCTAAAGTTGAACTTGGAGGCCAGGCATGTTGGCTCAGCCCTGTA 600
Db	193883	AGTGACAAATGGGTCTAAAGTTGAACTTGGAGGCCAGGCATGTTGGCTCAGCCCTGTA 193824
Qy	601	ATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTTGAGGCCAGGAGTTCGAGACCAG 660
Db	193823	ATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTTGAGGCCAGGAGTTCGAGACCAG 193764
Qy	661	CCTGGCCAAACATGTTGAAACCCCGTCTTACAAAAAAATACAAAAAATTAGCCGGGTG 720
Db	193763	CCTGGCCAAACATGTTGAAACCCCGTCTTACAAAAAATACAAAAAATTAGCCGGGTG 193704
Qy	721	GGTGATGGACACCTGTAGTACAGCTACTTGGGAGGCTGAGGAGGAGAAATTCCTGAAC 780
Db	193703	GGTGATGGACACCTGTAGTACAGCTACTTGGGAGGCTGAGGAGGAGAAATTCCTGAAC 193644
Qy	781	CCGGAGATGGAGGCTGCAGTGAAGTCAAGCTCAGGCCACTGCGCTCCAACTGGGCAACA 840
Db	193643	CTGGGAGATGGAGGCTGCAGTGAAGTCAAGCTCAGGCCACTGCGCTCCAACTGGGCAACA 193584
Qy	841	GAGTAAGACTCCATCTC-AAAAAAAAAAAAAGCTGGATTTGGAGTGAATTAATAACAT 899
Db	193583	GAGTAAGACTCCATCTC-AAAAAAAAAAAAAGCTGGATTTGGAGTGAATTAATAACAT 193524
Qy	900	TCTCCCTCTCTCTCTTTTGGCTGTGTCTCCATCTGTCTTTTTCGCAATTTCTTCATC 959
Db	193523	TCTCCCTCTCTCTCTTTTGGCTGTGTCTCCATCTGTCTTTTTCGCAATTTCTTCATC 193464
Qy	960	TCTGTACTTTCCATCTCTGTGTGTCTTCCATCTGCTTCTCCATGTATGGGCATCTCT 1019
Db	193463	TCTGTACTTTCCATCTCTGTGTGTCTTCCATCTGCTTCTCCATGTATGGGCATCTCT 193404

Qy	1020	GGGTCTCTCATGTCTCTTCTGCCCACACTTTGGCACATCTCTGCTCTCTCATGCCCCCT 1079
Db	193403	GGGTCTCTCATGTCTCTTCTGCCCACACTTTGGCACATCTCTGCTCTCTCATGCCCCCT 193344
Qy	1080	TTCTCTCTGCGAGGGTGAATCTGGGGGGGCTGTGGTGTGCAATGGCTCCCTCGAGGACT 1139
Db	193343	TTCTCTCTGCGAGGGTGAATCTGGGGGGGCTGTGGTGTGCAATGGCTCCCTCGAGGACT 193284
Qy	1140	CGTGTCTCTGGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCT 1199
Db	193283	CGTGTCTCTGGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCT 193224
Qy	1200	CTGCAAGTTTCAACCAAGTGGATCCAGGAAACCATCCAGGGCAACTCTCTGAGTCAATCCAGG 1259
Db	193223	CTGCAAGTTTCAACCAAGTGGATCCAGGAAACCATCCAGGGCAACTCTCTGAGTCAATCCAGG 193164
Qy	1260	ACTCAGACACCGGCATTCCTCCACCTGCTGCGAGGAGACGCTCTGACACTCTCTTCAGACCC 1319
Db	193163	ACTCAGACACCGGCATTCCTCCACCTGCTGCGAGGAGACGCTCTGACACTCTCTTCAGACCC 193104
Qy	1320	TCATTCCTTCCAGAGATGTTGAGAATGTTCACTCTCCAGGCCCTGACCCCATGTCTCC 1379
Db	193103	TCATTCCTTCCAGAGATGTTGAGAATGTTCACTCTCCAGGCCCTGACCCCATGTCTCC 193044
Qy	1380	TGACTCAGGGTCTGCTTCTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCG 1439
Db	193043	TGACTCAGGGTCTGCTTCTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCG 192984
Qy	1440	GAACAAATTTCCAAACATGTCAGGGGGGGTGGTCTCAATCTCTCTGCGGACATTTTC 1499
Db	192983	GAACAAATTTCCAAACATGTCAGGGGGGGTGGTCTCAATCTCTCTGCGGACATTTTC 192924
Qy	1500	ATCCTCAAGCTCAGGGCCCATCTCTCTGCGAGCTCTGACCCCAATTTAGTCCAGAAA 1559
Db	192923	ATCCTCAAGCTCAGGGCCCATCTCTCTGCGAGCTCTGACCCCAATTTAGTCCAGAAA 192864
Qy	1560	TAAACTGAGAAG 1571
Db	192863	TAAACTGAGAAG 192852

RESULT 8  
AC130782/c

AC130782	200792 bp	DNA	linear	HTG 22-NOV-2002
LOCUS	Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12 ordered pieces.			
DEFINITION				
ACCESSION	AC130782	GI:25167101		
VERSION	AC130782.2	HTGS PHASE2; HTGS DRAFT.		
KEYWORDS	Pan troglodytes (chimpanzee)			
SOURCE	Pan troglodytes			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.			
REFERENCE	1 (bases 1 to 200792)			
AUTHORS	Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lario, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Standrip, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.			
TITLE	NISC Comparative Sequencing Initiative			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 200792)			
AUTHORS	Green, E.D.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA			
REFERENCE	3 (bases 1 to 200792)			
AUTHORS	Green, E.D.			



TITLE Direct Submission  
 JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 COMMENT On Nov 22, 2002 this sequence version replaced gi:22218452.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)  
 ----- Project Information  
 Center project name: dhz  
 Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 196566 bases at least Q40  
Consensus quality: 197883 bases at least Q30  
Consensus quality: 198879 bases at least Q20  
Insert size: 215000; agarose-fp  
Insert size: 199692; sum-of-contigs  
Quality coverage: 9.03x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 28306: contig of 28306 bp in length  
 \* 28307 28406: gap of unknown length  
 \* 28407 37856: contig of 9450 bp in length  
 \* 37857 37956: gap of unknown length  
 \* 37958 73522: contig of 35566 bp in length  
 \* 73523 73622: gap of unknown length  
 \* 73623 83567: contig of 9945 bp in length  
 \* 83568 83667: gap of unknown length  
 \* 83668 88817: contig of 5150 bp in length  
 \* 88818 125611: contig of unknown length  
 \* 125612 125711: gap of unknown length  
 \* 125712 159879: contig of 34168 bp in length  
 \* 159880 159979: gap of unknown length  
 \* 159980 174698: contig of 14719 bp in length  
 \* 174699 186382: gap of unknown length  
 \* 186383 186482: contig of 11584 bp in length  
 \* 186483 193344: gap of unknown length  
 \* 193345 193444: contig of 6862 bp in length  
 \* 193445 199363: contig of 5919 bp in length  
 \* 199364 199463: gap of unknown length  
 \* 199464 200792: contig of 1329 bp in length

<b>FEATURES</b>	<b>SOURCE</b>
-----------------	---------------

	/db_xref="taxon:9598"	
	/clone="CH251-355A20"	
	/clone_lib="CH251"	
misc_feature	1..28306	
	/note="assembly_fragment"	
	clone_end:SP6	
	vector_side:left	
misc_feature	28407..37856	
	/note="assembly_fragment"	
misc_feature	37957..73522	
	/note="assembly_fragment"	
misc_feature	73623..83567	
	/note="assembly_fragment"	
misc_feature	83668..88817	
	/note="assembly_fragment"	
misc_feature	88918..125611	
	/note="assembly_fragment"	
misc_feature	125712..159879	
	/note="assembly_fragment"	
misc_feature	159980..174698	
	/note="assembly_fragment"	
misc_feature	174799..186382	
	/note="assembly_fragment"	
misc_feature	186483..193344	
	/note="assembly_fragment"	
misc_feature	193445..199363	
	/note="assembly_fragment"	
misc_feature	199464..200792	
	/note="assembly_fragment"	
	clone_end:T7	
	vector_side:right	
 ORIGIN		
Query Match 94.3%; Score 1482; DB 2; Length 200792;		
Best Local Similarity 98.1%; Pred. No. 0;		
Matches 1542; Conservative 0; Mismatches 25; Indels 5; Gaps 4;		
Qy	1 AGGGAGGTATGGGAATTGAAACAGGAAAACAAATTAGTCCAAAGCAATGGATTCTTA 60	
Dd	69290 AGGGAGGTATGGGAATTGAGACAGGAAAACAAATTAGTCCAAAGCAATGGATTCTTA 69231	
Qy	61 TTGGGAGTGATTCTGCCCTAGAAGCACTGGCAATACCAGGAGACATTTTGGTGTGTC 120	
Dd	69230 TTGGGAGTGATTCTGCCCTAGAAGCACTGGCAATACCAGGAGACATTTTGGTGTGTC 69171	
Qy	121 CAACTATATGGAGGGGCATTACTGCCAACTAATGGATAGATGCCAAGTGTCTGTTCAAC 180	
Dd	69170 CAACTATATGGAGGGGCATTACTGCCAACTAATGGATAGAGGCCAAGGTGCTGTTCAAC 69111	
Qy	181 ATGCTATGATGACACGGCAGGCCTCCACAACCAACCATTCAGCTTCAGATGCCCAC 240	
Dd	69110 GTGCTATGATGACACGGCAGGCCTCCACAACCAACCATTCAGCTTCAGATGCCCAC 69051	
Qy	241 AGTGCCCGATCGAGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTCGAGAAGGGAG 300	
Dd	69050 AGTGCCCGATCGAGAACCCCTCATCCAGGGGCTGAGAACCTGTATTTTCGAGAAGGGAG 68991	
Qy	301 GTATAAGTAGTGGTTGGTGGAGAAATGGGGAAGGAGGTGTGTGCCAGTAAGAAAAATA 360	
Dd	68990 GTATAAGTAGTGGTTGGTGGAGAAATGGGGAAGGAGGTGTGTGCCAGTAAGAAAAATA 68931	
Qy	361 GGCCTGCACAGGCTGGAGGGGAGATGAGAGAAAGGAGCGGAGAGATACACGATGA 420	
Dd	68930 GGCCTGCACAGGCTGGAGGGGAGATGAGAGAAAGGAGCGGAGAGATACACGATGA 68871	
Qy	421 GGAGACAGGCTGGAAACGAAAGTAGACACGAAGATTCCAGATGTGGAGAGAAAGGGTCA 480	
Dd	68870 GGAGACAGGCTGGAAACGAAAGTAGACACGNAGATTCCAGATGTGGAGAGAAAGGGTCA 68811	
Qy	481 CAGACCCCCCAAATGATGTGTGGACAAACAGGAATCTGGAAAGGAGGAAGATGGAGTGGAG 540	
Dd	68810 CAG-GCCCCCCAATGATGTGT-GACAAACAGGAATCTGGAAAGGAGGAAGATGGAGTGGAG 68753	





SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horridan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature gene sets  
JOURNAL Patent: WO 0194629-A 2333 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES Location/Qualifiers  
source 1..586  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 30.7%; Score 482; DB 6; Length 586;  
Best Local Similarity 100.0%; Pred. No. 3.8e-131; Mismatches 0; Indels 0; Gaps 0;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 1149  
DB 501 CAGGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 442  
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
DB 441 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 382  
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269  
DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA 322  
QY 1270 CCGCATCCCACTGTCTGCGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 1329  
DB 321 CCGCATCCCACTGTCTGCGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 442  
QY 1330 CCAGAGATGTGAGATGTTTCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269  
DB 381 CCAGAGATGTGAGATGTTTCAATCTCTGAGTCAATCCAGGACTCAGCACA 322  
QY 1270 CCGCATCCCACTGTCTGCGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 1329  
DB 321 CCGCATCCCACTGTCTGCGGGGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTGG 442  
QY 1330 CCAGAGATGTGAGATGTTTCAATCTCTGAGTCAATCCAGGACTCAGCACA 1269  
DB 261 CCAGAGATGTGAGATGTTTCAATCTCTGAGTCAATCCAGGACTCAGCACA 322  
QY 1390 GTCTGCTTCCCACTGTGGGCTGACCGTGTCTCTAGTGAACCCCTGGGAACTTC 1449  
DB 201 GTCTGCTTCCCACTGTGGGCTGACCGTGTCTCTAGTGAACCCCTGGGAACTTC 142  
QY 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1509  
DB 141 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 82  
QY 1510 TCAGGGCCCACTCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569  
DB 81 TCAGGGCCCACTCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 22  
QY 1570 AG 1571  
DB 21 AG 20  
RESULT 11  
G41906/c  
LOCUS  
DEFINITION SHGC-56840 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G41906  
VERSION G41906.1 GI:3668239  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 586)  
Myers, R.M.  
TITLE Human STSs (1998)

## JOURNAL COMMENT

Unpublished (1998)

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu  
Primer A: AGAGACACGGTCAGCCCAAT  
Primer B: GCCAACTCTGAGTCATCCC  
STS size: 188  
PCR Profile:  
Initial incubation: 95 degrees C for 10 minutes  
Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Amplitaq Gold Polymerase: 0.07 units/ul  
Total Vol: 5 ul

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Amplitaq Gold Polymerase: 0.07 units/ul

Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

Prepared with primer pairs derived from W73140 -- Unigene.

## FEATURES

Location/Qualifiers

1..586

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/map="19"

/clone\_lib="Human"

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

167..354

167..186

complement(335..354)

primer\_bind

primer\_bind

complement(335..354)

STS

Qy 1450 CAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509  
Db 141 CAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 82  
Qy 1510 TCAGGGCCCATCTCTCTGCGAGCTTGACCCAAATTTAGTCCAGAAATAACTGAGA 1569  
Db 81 TCAGGGCCCATCTCTCTGCGAGCTTGACCCAAATTTAGTCCAGAAATAACTGAGA 22  
Qy 1570 AG 1571  
Db 21 AG 20  
RESULT 12  
AX429955/c  
LOCUS  
DEFINITION Sequence 94 from Patent WO0198339.  
ACCESSION AX429955  
VERSION AX429955.1 GI:21541119  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Reed, S.G., Xu, J., Dillon, D.C., Retter, M.W. and Harlocker, S.L.  
TITLE Compositions and methods for the therapy and diagnosis of breast cancer  
JOURNAL Patent: WO 0198339-A 94 27-DEC-2001;  
CORIXA CORPORATION (US)  
FEATURES  
source 1..735  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 30.7%; Score 482; DB 6; Length 735;  
Best Local Similarity 100.0%; Pred. No. 3.9e-131;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1090 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1149  
Db 500 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 441  
Qy 1150 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
Db 440 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 381  
Qy 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269  
Db 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 321  
Qy 1270 CCGGCATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1329  
Db 320 CCGGCATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 261  
Qy 1330 CCAGAGATTGAGATTGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTC 1389  
Db 260 CCAGAGATTGAGATTGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTC 201  
Qy 1390 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
Db 200 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 141  
Qy 1450 CAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509  
Db 140 CAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 81  
Qy 1510 TCAGGGCCCATCTCTCTGCGAGCTTGACCCAAATTTAGTCCAGAAATAACTGAGA 1569  
Db 80 TCAGGGCCCATCTCTCTGCGAGCTTGACCCAAATTTAGTCCAGAAATAACTGAGA 21

Qy 1570 AG 1571  
Db 20 AG 19  
RESULT 13  
BD139877/c  
LOCUS  
DEFINITION 738 bp DNA linear PAT 18-SEP-2002  
Compounds for immunotherapy and diagnosis of breast cancer and methods for their use.  
ACCESSION BD139877  
VERSION BD139877.1 GI:23234822  
KEYWORDS JP 2002507387-A/94.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Reed, S.G. and Xu, J.  
TITLE Comounds for immunotherapy and diagnosis of breast cancer and methods for their use  
JOURNAL Patent: JP 2002507387-A 94 12-MAR-2002;  
CORIXA CORP  
COMMENT OS Homo sapiens (human)  
PN JP 2002507387-A/94  
PD 12-MAR-2002  
PF 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253, 24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627, 17-JUL-1998 US 09/118554 PI STEVEN  
G REED, JIANGCHUN XU  
PC C12N15/09, A61K38/00, A61K39/00, A61K39/39, A61K39/395, A61P35/00,  
PC C07K14/47,  
PC  
C07K16/18, C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/ PC  
53,  
PC G01N33/577, C12N15/00, A61K37/02, C12N5/00  
CC Compounds for immunotherapy and diagnosis  
of breast cancer and  
CC methods for  
CC their use  
FH Key Location/Qualifiers  
FT source 1..738  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
FEATURES  
source 1..738  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 30.7%; Score 482; DB 6; Length 738;  
Best Local Similarity 100.0%; Pred. No. 3.9e-131;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1090 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 1149  
Db 503 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCTGG 444  
Qy 1150 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
Db 443 GGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 384  
Qy 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269  
Db 383 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 324  
Qy 1270 CCGGCATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1329  
Db 323 CCGGCATCCCACTGCTGAGGAGACCGCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 264  
Qy 1330 CCAGAGATTGAGATTGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTC 1389

263 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCAATGTCCTCTGGACTCAGG 204  
QY GTCTGCTTCCCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1449  
Db GTCTGCTTCCCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTC 144  
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACATTTTCATCTCAAGC 1509  
Db 143 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACATTTTCATCTCAAGC 84  
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569  
Db 83 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 24  
QY 1570 AG 1571  
Db 23 AG 22  
RESULT 14  
LOCUS CQ720600 1143 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 6534 from Patent WO02068579.  
ACCESSION CQ720600  
VERSION CQ720600.1 GI:42281457  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
PATENT: WO 02068579-A 6534 06-SEP-2002;  
PE CORPORATION (NY) (US)  
FEATURES  
source  
1. .1143  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 30.7%; Score 482; DB 6; Length 1143;  
Best Local Similarity 100.0%; Pred. No. 4e-131;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGATTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 1149  
Db 657 CAGGGTGATTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 716  
QY 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
Db 717 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 776  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCCACTCTCTAGTTCATCCAGGACTCAGCACA 1269  
Db 777 ACCAAGTGGATCCAGGAAACCATCCAGGCCCACTCTCTAGTTCATCCAGGACTCAGCACA 836  
QY 1270 CCGGCATCCCACTGTGTGAGGACAGCCCTGACACTCTCTTTGAGACCCCTCATTCCTTC 1329  
Db 837 CCGGCATCCCACTGTGTGAGGACAGCCCTGACACTCTCTTTGAGACCCCTCATTCCTTC 896  
QY 1330 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 1389  
Db 897 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 956  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1449  
Db 957 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1016  
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCCCTGGGGCACATTTTCATCTCAAGC 1509

1017 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCTCCCTGGGACATTTTCATCTCAAGC 1076  
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569  
Db 1077 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1136  
QY 1570 AG 1571  
Db 1137 AG 1138  
RESULT 15  
LOCUS AY461805 1301 bp mRNA linear PRI 03-DEC-2003  
DEFINITION Homo sapiens kallikrein 5 isoform 3 preproprotein (KLK5) mRNA,  
complete cds, alternatively spliced.  
ACCESSION AY461805  
VERSION AY461805.1 GI:38564772  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL Michael, I.P., Kurlender, L., Du, D.C. and Diamandis, E.P.  
Cloning of new splice variants of the human kallikrein gene 5  
Unpublished  
REFERENCE 2 (bases 1 to 1301)  
AUTHORS Michael, I.P., Kurlender, L., Du, D.C. and Diamandis, E.P.  
TITLE Direct Submission  
JOURNAL Submitted (10-NOV-2003) Pathology and Laboratory Medicine, Mount  
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
Canada  
FEATURES  
source  
1. .1301  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.4"  
1. .1301  
/gene="KLK5"  
12. .626  
/genes="KLK5"  
/notes="alternatively spliced"  
/codon\_start=1  
/product="kallikrein 5 isoform 3 preproprotein"  
/protein\_id="AA23814.1"  
/db\_xref="GI:38564772"  
/translation="MATARPMMVLCALITALLGVTEHVLANNNDVSCDHPNTPVS  
GNSQDILGAGAGEDARDSDSSRIINGSDCDMHTQWQAALLRPNQLYCGAVLVHPQW  
LITAAHCRKVRVRLGHVLSLSPVYESGQMGQFQVKSIPHPGYSHPGHSNDLMLIKLN  
RRIRPTKDVRLPINVSSHCPSPAGTKCLVSGWGTTKSPQGECPGSS"  
ORIGIN  
Query Match 30.7%; Score 482; DB 9; Length 1301;  
Best Local Similarity 100.0%; Pred. No. 4.1e-131;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGATTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 1149  
Db 820 CAGGGTGATTTCTGGGGGCGCTGTGCTGCAATGGCTCCCTGCAGGACTCGTGTCTCGG 879  
QY 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
Db 880 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 939  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCCACTCTCTAGTTCATCCAGGACTCAGCACA 1269  
Db 940 ACCAAGTGGATCCAGGAAACCATCCAGGCCCACTCTCTAGTTCATCCAGGACTCAGCACA 999  
QY 1270 CCGGCATCCCACTGTGTGAGGACAGCCCTGACACTCTCTTTGAGACCCCTCATTCCTTC 1329

```
Db      ||||| 1000 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTGACAGCCCTCATTCCTTC 1059
Qy      ||||| 1330 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCCTGGACTCAGG 1389
Db      ||||| 1060 CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCCTGGACTCAGG 1119
Qy      ||||| 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTTC 1449
Db      ||||| 1120 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTTC 1179
Qy      ||||| 1450 CAAAACGTCTCAGGCGGGGGTTGCGTCTCAATCTCCTGGGGCACTTTCATCCTCAAGC 1509
Db      ||||| 1180 CAAAACGTCTCAGGCGGGGGTTGCGTCTCAATCTCCTGGGGCACTTTCATCCTCAAGC 1239
Qy      ||||| 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1569
Db      ||||| 1240 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAATAAACTGAGA 1299
Qy      ||||| 1570 AG 1571
Db      ||||| 1300 AG 1301
```

Search completed: July 4, 2005, 09:55:43  
Job time : 4534.17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 552.864 Seconds  
(without alignments)  
16821.327 Million cell updates/sec

Title: US-09-936-271c-13\_COPY\_10000\_11570  
Perfect score: 1571  
Sequence: 1 agggagtgatgggaattga.....ccagaaataaactgagaag 1571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Geneseqn16Dec04:\*  
2: Geneseqn1980s:\*  
3: Geneseqn1980s:\*  
4: Geneseqn2000as:\*  
5: Geneseqn2001as:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	11570	3	AAA95905
2	1571	100.0	11570	12	AdK52482 Human KLK
3	1571	100.0	11570	13	AdR72623 Human ren
4	1571	100.0	11570	13	AdR72875 Human ova
5	541	34.4	567	12	ACH74409 Human gen
C 6	482	30.7	586	6	ABL63579 Breast ca
C 7	482	30.7	586	6	ABL63996 Breast ca
C 8	482	30.7	735	3	AACT9469
C 9	482	30.7	735	6	ABK29013
C 10	482	30.7	738	2	AAK84240
C 11	482	30.7	1260	10	ADRB0489
C 12	482	30.7	1260	11	ADN39197
C 13	482	30.7	1260	13	AdR32550 Breast ca
C 14	482	30.7	1302	2	AAK57990
C 15	482	30.7	1302	2	AAK57989
C 16	482	30.7	1381	2	AAZ06259
C 17	482	30.7	1381	8	ADA40370 Human sec
C 18	482	30.7	1381	10	ADA56535 Gene enco
C 19	482	30.7	1387	13	AdR72622 Human ren
C 20	482	30.7	1387	13	AdR72874 Human ova

21	482	30.7	1439	2	AAZ06260
22	482	30.7	1439	8	ADA40371
23	482	30.7	1439	10	ADA56536 Gene enco
24	482	30.7	1441	8	ADA40368
25	482	30.7	1441	10	ADA56533 Gene enco
26	482	30.7	1481	12	ADQ22212
27	482	30.7	1490	8	ADA40367
28	482	30.7	1490	10	ADA56532
29	482	30.7	1499	2	AAK60578
30	482	30.7	1516	2	AAZ06245
31	482	30.7	1516	8	ADA40369
32	482	30.7	1516	10	ADA56534
33	482	30.7	1570	3	AAZ65070
34	482	30.7	1570	4	AAK21471
35	482	30.7	1570	5	AAF44216
36	482	30.7	1570	6	ABK28605
37	482	30.7	1570	8	ACA64363
38	482	30.7	1570	8	ACA03830
39	482	30.7	1570	8	ABX89368
40	482	30.7	1570	8	ABX80822
41	482	30.7	1570	8	ACD44331
42	482	30.7	1570	8	ACD42022
43	482	30.7	1570	8	ABX79502
44	482	30.7	1570	8	ACA93523
45	482	30.7	1570	8	ABX81205

ALIGNMENTS

RESULT 1  
AAA95905  
ID AAA95905 standard; DNA; 11570 BP.  
XX  
AC AAA95905;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KLK-L2 gene.  
XX  
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
KW kallikrein-like protein; serine protease; cytosolic; cancer;  
KW prostrate cancer; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-CA000258.  
XX  
PR 11-MAR-1999; 99US-0124260P.  
PR 01-APR-1999; 99US-0127386P.  
PR 21-JUL-1999; 99US-0144919P.  
XX  
PA (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
PI Yousef GM, Diamandis EP;  
XX  
DR WPI; 2000-587440/55.  
DR P-PSDB; AAB21296.  
XX  
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
XX protein mediated disorders, especially cancer.  
PS Claim 1; Page 143-149; 184pp; English.  
XX  
CC The present sequence is the coding sequence of the human KLK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyse the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding

CC	kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment,
CC	monitoring and diagnosis of cancers, especially prostate cancer. They
CC	can also be used to identify a substance that can associate with or
CC	mediate the biological activity of the proteins. Antibodies can be used
CC	to treat conditions mediated by the kallikrein-like proteins
XX	
SQ	Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;
	Query Match 100.0%; Score 1571; DB 3; Length 11570;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AGGAGGATATGGGGAATTGAAGACAGAGAAACAAAATAATAGTCCAAGCGAATGCATTCTTA 60
Db	
Db	10000 AGGAGGATATGGGGAATTGAAGACAGAGAAACCAAATTTAGTCCAGCGAATGCATTCTTA 10059
Qy	61 TTGGGAGTGATCTTGCCCTTAGAAGACACTGGCAATACAGAGAGACATTTTTGTGTGCA 120
Db	
Db	10060 TTGGGAGTGATCTTGCCCTTAGAAGACACTGGCAATACCAGGAGACATTTTTGTGTGCA 10119
Qy	121 CAACTATATGGGGGCGATTACTGCCAACTAATGGATAGATGCCAAGTGTGCTGTTCAAC 180
Db	
Db	10120 CAACTATATGGGGGCGATTACTGGCAACTAATGGATAGATGCCAAGTGTGCTGTTCAAC 10179
Qy	181 ATGCTATGATGCACAGGCAGGCCTCCAAACAACCAATTATTCAGCTTCAGATCCCCAC 240
Db	
Db	10180 ATGCTATGATGCACAGGCAGGCCTCCAAACAACCAATTATTCAGCTTCAGATCCCCAC 10239
Qy	241 AGTGCCCAGATCAGAGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTGCGAAGAGGAG 300
Db	
Db	10240 AGTGCCCAGATCAGAGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTGCGAAGAGGAG 10299
Qy	301 GTATAGGATGGGTTGGTGGAGAAATGGGNAGGAGGTGTGTGTCAGTAAAGAGAATAA 360
Db	
Db	10300 GTATAGGATGGGTTGGTGGAGAAATGGGAAGGAAGGTGTGTGTCAGTAAAGAGAATAA 10359
Qy	361 GGCTCGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAGGAGGCGGAGAGATACAGATGA 420
Db	
Db	10360 GGCTCGCACAGGCTGGAGGGGAGAGTGAGANGAAGGAGGCGGNGAGATACAGANTGA 10419
Qy	421 GGGAGACAGGCTGGAAACAGAAAGTATAGACACGAAGATTCGAGATGTGGAGAGGAAGGGTCA 480
Db	
Db	10420 GGGAGACAGGCTGGAAACAGAAAGTATAGACACGAAGATTCGAGATGTGGAGAGGAAGGGTCA 10479
Qy	481 CAGACCCCCCGAAATGATGTGTGGACACACAGGAATCTCGGAAGAGGAAGATGGAGTGGAG 540
Db	
Db	10480 CAGACCCCCCGAAATGATGTGTGGACCAACAGGAATCTCGGAAGAGGAAGATGGAGTGGAG 10539
Qy	541 AGTGACAAATGGGGTCTTAAAGGTTTGAATTTGGAGGCCAGGCATGTGGCTCACGGCTGTA 600
Db	
Db	10540 AGTGACAAATGGGGTCTTAAAGGTTTGAATTTGGAGGCCAGGCATGTGGCTCACGGCTGTA 10599
Qy	601 ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTTGAGGCCAGGAGTTTCGAGACCAG 660
Db	
Db	10600 ATCCCAACACTTTGGAGGCTGAGGTGGCGGAATCACTTTGAGGCCAGGAGTTTCGAGACCAG 10659
Qy	661 CCTGGCCAAACATGGTGAACCCCGTCTTACAAAAAATAACAAAAATTAGCCGGGTGT 720
Db	
Db	10660 CCTGGCCAAACATGGTGAACCCCGTCTTACAAAAAATAACAAAAATTAGCCGGGTGT 10719
Qy	721 GGTGATGGACACCTGTAGTCACAGCTACTTTGGGAGGCTCAGCGCAGGAGAAATGCTTTGAAC 780
Db	
Db	10720 GGTGATGGACACCTGTAGTCACAGCTACTTTGGGAGGCTCAGCGCAGGAGAAATGCTTTGAAC 10779
Qy	781 CCGGGAGATGGAGGCTGCAGTGAGCTGAGTCAAGGCCACTTCGGCTCCAACTCGGGCAACA 840
Db	
Db	10780 CCGGGAGATGGAGGCTGCAGTGAGCTGAGTCAAGGCCACTTCGGCTCCAACTCGGGCAACA 10839
Qy	841 GAGTAAGACTCCATCTCAAAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATTAACATT 900
Db	
Db	10840 GAGTAAGACTCCATCTCAAAAAAAAAAAGCTGGATTTGGAGTGAATAATTAATTAACATT 10899

XX Diamandis EP;  
XX WPI; 2004-239232/22.  
DR P-PSDB; ADK52481.  
XX  
XX Detecting kallikrein 5 associated with cancer, useful in diagnosing,  
PT monitoring, detecting, imaging and treating breast or ovarian carcinoma,  
PT comprises comparing the detected amount of kallikrein 5 in a sample with  
PT a standard sample.  
XX  
XX Disclosure; SEQ ID NO 2; 59pp; English.  
XX  
XX The present invention relates to detecting kallikrein 5 associated with  
CC breast or ovarian cancer in a patient comprises detecting in the sample  
CC kallikrein 5 and comparing the detected amount with an amount detected  
CC for a standard. The method is useful in detecting kallikrein 5 associated  
CC with breast or ovarian cancer in a patient. The methods and kits are  
CC useful in diagnosing, monitoring, detecting, imaging and treating breast  
CC or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for  
CC preventing and treating breast and ovarian cancer and for stimulating or  
CC enhancing antibody production or for inducing an immune response. The  
CC present sequence represents human kallikrein 5 encoding sequence.  
XX  
XX Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;  
SQ

Query Match 100.0%; Score 1571; DB 12; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGAGGTATGGGGAATTGAAGACAGGAAACACAAATTTAGTCCAAAGCGAATGGATTCTTA 60  
DB 10000 AGGGAGGTATGGGGAATTGAAGACAGGAAACACAAATTTAGTCCAAAGCGAATGGATTCTTA 10059  
QY 61 TTGGAGGTGATTCTGCCCTTAGAAGACACTGGCAATACACAGGAGACATTTTGTGTGCA 120  
DB 10060 TTGGAGGTGATTCTGCCCTTAGAAGACACTGGCAATACACAGGAGACATTTTGTGTGCA 10119  
QY 121 CAATATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGCTGTTCAAC 180  
DB 10120 CAATATATGAGGGGCAATTACTGGCAACTAATGGATAGATGCCAAGTGCTGTTCAAC 10179  
QY 181 ATGTATGATGACACGGGAGCCCTTCAATCCAGGAGGCTGAGAACCGTATTTTGCAGAAAGGGAG 240  
DB 10180 ATGTATGATGACACGGGAGCCCTTCAATCCAGGAGGCTGAGAACCGTATTTTGCAGAAAGGGAG 10239  
QY 241 AGTCCCAAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTGCAGAAAGGGAG 300  
DB 10240 AGTCCCAAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTGCAGAAAGGGAG 10299  
QY 301 GTATAAGGATGGTGTGGAGAAATGGGGAAGGAGTGTGTGTCAGTAAGAAATAA 360  
DB 10300 GTATAAGGATGGTGTGGAGAAATGGGGAAGGAGTGTGTGTCAGTAAGAAATAA 10359  
QY 361 GGCTTCACAGCTGGAGGGAGAGTGGAGAGAAAGGAGCGGAGGATACAGATGA 420  
DB 10360 GGCTTCACAGCTGGAGGGAGAGTGGAGAGAAAGGAGCGGAGGATACAGATGA 10419  
QY 421 GGGAGACAGGCTGGAAACAGAAAGTAGACCAAGATTCCAGATGTGGAGAGAAAGGGTCA 480  
DB 10420 GGGAGACAGGCTGGAAACAGAAAGTAGACCAAGATTCCAGATGTGGAGAGAAAGGGTCA 10479  
QY 481 CAGACCCCGGAAATGATGTGGACAAACAGGAATCTGGAAGAGGAGATGGAGTGGAG 540  
DB 10480 CAGACCCCGGAAATGATGTGGACAAACAGGAATCTGGAAGAGGAGATGGAGTGGAG 10539  
QY 541 AGTGACAAATGGGTCTAAAGGTTGAACCTTGGAGCCAGGCAATGGTGGCTCACGCCGTGA 600  
DB 10540 AGTGACAAATGGGTCTAAAGGTTGAACCTTGGAGCCAGGCAATGGTGGCTCACGCCGTGA 10599  
QY 601 ATCCCAACACTTTGGAGGCTGAGTGGCGGAATCACTTTGAGGCCAGGAGTTTCGAGACCAG 660  
DB 10600 ATCCCAACACTTTGGAGGCTGAGTGGCGGAATCACTTTGAGGCCAGGAGTTTCGAGACCAG 10659

## RESULT 3

AD72623  
ID AD72623 standard; DNA; 11570 BP.  
XX  
XX AC AD72623;  
XX DT 02-DEC-2004 (first entry)

QY 661 CCTGGCCACATCGTGAACCCCGTCTCTACAAAAAATAACAAAAATTAGCCGGGTGT 720  
DB 10660 CCTGGCCACATCGTGAACCCCGTCTCTACAAAAAATAACAAAAATTAGCCGGGTGT 10719  
QY 721 GGTGATGGACACCTGTAGTACACAGCTACTTTGGGAGGCTGAGGAGGAGAAATTTGTTCAAC 780  
DB 10720 GGTGATGGACACCTGTAGTACACAGCTACTTTGGGAGGCTGAGGAGGAGAAATTTGTTCAAC 10779  
QY 781 CCGGAGATGGAGGCTGACGTGAGCTGAGTCAAGGCCACTTGGCGTCCAACTGGCGCAACA 840  
DB 10780 CCGGAGATGGAGGCTGACGTGAGCTGAGTCAAGGCCACTTGGCGTCCAACTGGCGCAACA 10839  
QY 841 GAGTAGACTCCATCTCAAAAAAAGCTGATTTGGAGTGAATATTAATAACATT 900  
DB 10840 GAGTAGACTCCATCTCAAAAAAAGCTGATTTGGAGTGAATATTAATAACATT 10899  
QY 901 CTCCCTCTCTCTCTTTTGGCTGTCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT 960  
DB 10900 CTCCCTCTCTCTCTTTTGGCTGTCTCCATCTCTGTCTTTTCTGCAATTTCTTCATCT 10959  
QY 961 CTGTACTTTCCATCTCTGTGTGTCTTCCATCTGTCTTCCATCTATGGGATCTCTG 1020  
DB 10960 CTGTACTTTCCATCTCTGTGTGTCTTCCATCTGTCTTCCATCTATGGGATCTCTG 11019  
QY 1021 GGTCTCTCATGTCTCTTCTTCCCACTTTTGGCCACATCTCTGCCCTCTCTCATGCCCTTT 1080  
DB 11020 GGTCTCTCATGTCTCTTCTTCCCACTTTTGGCCACATCTCTGCCCTCTCTCATGCCCTTT 11079  
QY 1081 TCTCTCTCGCAGGCTGATTCTGGGGGCTGTGTGTCTGCAATGGCTCTCTGAGGAGCTC 1140  
DB 11080 TCTCTCTCGCAGGCTGATTCTGGGGGCTGTGTGTCTGCAATGGCTCTCTGAGGAGCTC 11139  
QY 1141 GTGTCTGGGAGATTACCTTTGTCCCGGCCCAACAGACCGGGTGTCTACAGAACCTC 1200  
DB 11140 GTGTCTGGGAGATTACCTTTGTCCCGGCCCAACAGACCGGGTGTCTACAGAACCTC 11199  
QY 1201 TGCAAGTTTCAACAAAGTGATCCAGGAACCACTCCAGGCCAACTCTGAGTCAATCCAGGA 1260  
DB 11200 TGCAAGTTTCAACAAAGTGATCCAGGAACCACTCCAGGCCAACTCTGAGTCAATCCAGGA 11259  
QY 1261 CTGACACACCGGCATCCCACTCTGTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCT 1320  
DB 11260 CTGACACACCGGCATCCCACTCTGTGAGGAGACAGCCCTGACACTCTCTTTCAGACCCCT 11319  
QY 1321 CATTTCTTCCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCT 1380  
DB 11320 CATTTCTTCCAGAGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCTCT 11379  
QY 1381 GGACTCAGGGTCTGTCTTCCCACTATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG 1440  
DB 11380 GGACTCAGGGTCTGTCTTCCCACTATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG 11439  
QY 1441 AACAAATTTCCAAACTGTCCAGGGCGGGGTGCGTCTCAATCTCCCTGGGGCACTTTTCA 1500  
DB 11440 AACAAATTTCCAAACTGTCCAGGGCGGGGTGCGTCTCAATCTCCCTGGGGCACTTTTCA 11499  
QY 1501 TCCTCAAGCTCAGGGCCCATCTCTTCTGTGAGCTGTGAGCCCAAAATTTAGTCCAGAAAT 1560  
DB 11500 TCCTCAAGCTCAGGGCCCATCTCTTCTGTGAGCTGTGAGCCCAAAATTTAGTCCAGAAAT 11559  
QY 1561 AAATCTGAGAAG 1571  
DB 11560 AAATCTGAGAAG 11570

XX Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.  
DE kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;  
KW cancer metastasis; chemotherapy; human; serine protease;  
KW chromosome 19q13.4; KLK5; ds; gene.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH 2221..11247  
FT /\*tag= b  
FT /product= "Human renal cell carcinoma-related kallikrein  
FT exon 5 (hK5) protein"  
FT 2221..2293  
FT /\*tag= a  
FT /number= 1  
FT intron 2294..4761  
FT /\*tag= c  
FT /number= 1  
FT exon 4762..5023  
FT /\*tag= d  
FT /number= 2  
FT intron 5024..5762  
FT /\*tag= e  
FT /number= 2  
FT exon 5763..6019  
FT /\*tag= f  
FT /number= 3  
FT intron 6020..6104  
FT /\*tag= g  
FT /number= 3  
FT exon 6105..6238  
FT /\*tag= h  
FT /number= 4  
FT exon 6239..11091  
FT /\*tag= i  
FT /number= 4  
FT exon 11092..11247  
FT /\*tag= j  
FT /number= 5  
XX WO2004077060-A2.  
XX 10-SEP-2004.  
XX 26-FEB-2004; 2004WO-CA000280.  
XX 27-FEB-2003; 2003US-0451382P.  
XX (MOUN ) MOUNT SINAI HOSPITAL.  
XX Diamandis EP, Petraki CD;  
XX WPI; 2004-662077/64.  
XX P-PSDB; ADR72621.  
XX Detecting kallikrein polypeptides/polynucleotides associated with renal  
XX cell carcinoma in a patient, for diagnosing/treating the disease,  
XX comprises detecting /identifying kallikrein polypeptides/polynucleotides  
XX in a sample.  
XX Example 1; SEQ ID NO 3; 53pp; English.  
XX The invention relates to a novel method for detecting kallikrein  
XX polypeptides, or the polynucleotides encoding them, associated with renal  
XX cell carcinoma. The method comprises obtaining a sample from a patient  
XX and detecting kallikrein polypeptides, or their encoding polynucleotides,  
XX where the kallikrein polypeptides are selected from the group consisting  
XX of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The  
XX detected amounts of the kallikrein polypeptides are compared to standard  
XX amounts. The molecules of the invention demonstrate cytostatic activity  
XX whilst the methods and kit may be useful for detecting, characterising,

CC preventing and treating renal cell carcinoma. Furthermore, the methods  
CC may be useful for evaluating the probability of the presence of malignant  
CC or pre-malignant cells and for detecting and quantitating tumour growth  
CC and cancer metastasis. Finally, the methods may be utilised to confirm  
CC the absence or removal of all tumour tissue following surgery, cancer  
CC chemotherapy and/or radiation therapy and to monitor cancer chemotherapy  
CC and tumour reappearance. The current sequence is that of the human  
CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the  
CC invention which encodes a secreted serine protease and is located at  
CC chromosome 19q13.4.  
XX

SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1571; DB 13; Length 11570;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAAATTAGTCCAAGCGAATGGATTCTTA 60

Db 10000 AGGGAGGTATGGGAAATTGAAGACAGGAAACACAAAATTAGTCCAAGCGAATGGATTCTTA 10059

Qy 61 TTGGGAGTGAATCTGCCCTTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA 120

Db 10060 TTGGGAGTGAATCTGCCCTTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA 10119

Qy 121 CAACTATATGGAGGGCATTTACTGGCAACTAATGGATAGATGCCAAGTGTCTTCAAC 180

Db 10120 CAACTATATGGAGGGCATTTACTGGCAACTAATGGATAGATGCCAAGTGTCTTCAAC 10179

Qy 181 ATGCTATATGCACACGGCAGGCTCCACAACAAACCATTTATCCAGTTCAGATGCCAC 240

Db 10180 ATGCTATATGCACACGGCAGGCTCCACAACAAACCATTTATCCAGTTCAGATGCCAC 10239

Qy 241 AGTGCCCATCGAGGAACCTCATCCAGGGGTGAGAACCGTATTTTTCAGAGGGGAG 300

Db 10240 AGTGCCCATCGAGGAACCTCATCCAGGGGTGAGAACCGTATTTTTCAGAGGGGAG 10299

Qy 301 GTATAAGGATGGTGTGGAGAAATGGGAAGAAAGTGTGTCTCAGTAAGAGAAATAA 360

Db 10300 GTATAAGGATGGTGTGGAGAAATGGGAAGAAAGTGTGTCTCAGTAAGAGAAATAA 10359

Qy 361 GGCCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAAGGAGCGGAGAGATACACGATGA 420

Db 10360 GGCCTGCACAGGCTGGAGGGGAGAGTGAGAGAGAAAAGGAGCGGAGAGATACACGATGA 10419

Qy 421 GGGAGACAGGCTGGAAACAGAAAGTAGACAGCAAGATTTCAGATCTCGAGAGAGGGTCA 480

Db 10420 GGGAGACAGGCTGGAAACAGAAAGTAGACAGCAAGATTTCAGATCTCGAGAGAGGGTCA 10479

Qy 481 CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 540

Db 10480 CAGACCCCCCGAAATGATGTGTGGACAAACAGGAATCTGGAAGAGGAAGATGGAGTGGAG 10539

Qy 541 AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 600

Db 10540 AGTGACAAATGGGGTCTAAAGTTGAACCTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 10599

Qy 601 ATCCCAACACTTTGGAGGCTGAGGTGGGGCAATCATTGAGGCCAGGAGTTCAGACCAG 660

Db 10600 ATCCCAACACTTTGGAGGCTGAGGTGGGGCAATCATTGAGGCCAGGAGTTCAGACCAG 10659

Qy 661 CTGGGCCAATGGTGAACCCCGTCTCTACAAAAAATAATACAAAAATTAGCGGGTGT 720

Db 10660 CTGGGCCAATGGTGAACCCCGTCTCTACAAAAAATAATACAAAAATTAGCGGGTGT 10719

Qy 721 GGTGATGGACACCTGTAGTCACAGTACTTTGGAGGCTGAGGAGGAGAAATGCTTGAAC 780

Db 10720 GGTGATGGACACCTGTAGTCACAGTACTTTGGAGGCTGAGGAGGAGAAATGCTTGAAC 10779

Qy 781 CCGGAGATGGAGGCTGCAGTGAAGTCAAGGCTAGGCTCCAACTCCAACTGGGCAACA 840

Db 10780 CCGGAGATGGAGGCTGCAGTGAAGTCAAGGCTAGGCTCCAACTCCAACTGGGCAACA 10839





Db 10000 AGGAGGATATGGGGAATTGAGACAGGAAAACACAAATTAGTCCAGCGAATGGAATTTCTA 10059  
Qy 61 TTGGAGGTGATTTCTGCCCCCTAGAGACACTGGCAATACACAGGAGACATTTTGGTGTCTCA 120  
Db 10060 TTGGAGGTGATTTCTGCCCCCTAGAGACACTGGCAATACACAGGAGACATTTTGGTGTCTCA 10119  
Qy 121 CAATATATGAGGGGGCAATTAATGCGCAACTAATGATAGATGCGCAAGTGTGCTTTCAAC 180  
Db 10120 CAATATATGAGGGGGCAATTAATGCGCAACTAATGATAGATGCGCAAGTGTGCTTTCAAC 10179  
Qy 181 ATGCTATGATGCAACACGGAGGCTTCCACAAACCAATATCCAGTTCAGATGCCCCAC 240  
Db 10180 ATGCTATGATGCAACACGGAGGCTTCCACAAACCAATATCCAGTTCAGATGCCCCAC 10239  
Qy 241 AGTGCCACAGATCGAGGAACCTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 300  
Db 10240 AGTGCCACAGATCGAGGAACCTCATCCAGGGCTGAGAACCGTATTTTTCAGAGGGAG 10299  
Qy 301 GTATAAGGATGGGTGCTGAGAAATGGGAAGGAAGTGTGTGTCAGTAAGAGAAATAA 360  
Db 10300 GTATAAGGATGGGTGCTGAGAAATGGGAAGGAAGTGTGTGTCAGTAAGAGAAATAA 10359  
Qy 361 GGCTGTGACAGGCTGAGGGGAGAGTGAAGAGAAAGGAGGGGAGAGATACACGATGA 420  
Db 10360 GGCTGTGACAGGCTGAGGGGAGAGTGAAGAGAAAGGAGGGGAGAGATACACGATGA 10419  
Qy 421 GGGAGACAGGCTGAGAAACAGAAAGTAGAGACGAGATTCGAGATGTCAGAGGAGGTC 480  
Db 10420 GGGAGACAGGCTGAGAAACAGAAAGTAGAGACGAGATTCGAGATGTCAGAGGAGGTC 10479  
Qy 481 CAGACCCCCCGAAATGATGTGCGAACACAGGAATCTGGAAGAGGAAGATGGAAGTGGAG 540  
Db 10480 CAGACCCCCCGAAATGATGTGCGAACACAGGAATCTGGAAGAGGAAGATGGAAGTGGAG 10539  
Qy 541 AGTGACAAATGGGTCTAAAGTTGAACTTGAGGGCCAGGCAATGGTGGCTCAGCCCTGTA 600  
Db 10540 AGTGACAAATGGGTCTAAAGTTGAACTTGAGGGCCAGGCAATGGTGGCTCAGCCCTGTA 10599  
Qy 601 ATCCACACATTTGGAGGCTGAGGTGGCGCAATCATTGAGGCCAGGAGTTCGAGACCAG 660  
Db 10600 ATCCACACATTTGGAGGCTGAGGTGGCGCAATCATTGAGGCCAGGAGTTCGAGACCAG 10659  
Qy 661 CCTGCCCCAATGCTGAAACCCCTCTCTACAAAAAATAACAAAAAATAGCCGGGTGT 720  
Db 10660 CCTGCCCCAATGCTGAAACCCCTCTCTACAAAAAATAACAAAAAATAGCCGGGTGT 10719  
Qy 721 GGTGATGGACACCTGTAGTACAGCTACTGGGAGGCTGAGGCGAGGAAATTCCTGAAC 780  
Db 10720 GGTGATGGACACCTGTAGTACAGCTACTGGGAGGCTGAGGCGAGGAAATTCCTGAAC 10779  
Qy 781 CCGGGAGATGGAGCTGCAGTGCAGTGCAGGTCAGGCGCACCTGCGCTCCAACTGGGCAACA 840  
Db 10780 CCGGGAGATGGAGCTGCAGTGCAGTGCAGGTCAGGCGCACCTGCGCTCCAACTGGGCAACA 10839  
Qy 841 GAGTAAGACTCCATCTCAAAAAAATAAAGCTGGAATTTGGAGTGAATAATTAATAACATT 900  
Db 10840 GAGTAAGACTCCATCTCAAAAAAATAAAGCTGGAATTTGGAGTGAATAATTAATAACATT 10899  
Qy 901 CTCCT 960  
Db 10900 CTCCT 10959  
Qy 961 CTGTACTTTCCATCTCTGT 1020  
Db 10960 CTGTACTTTCCATCTCTGT 11019  
Qy 1021 GGTCTCTCATGCT 1080  
Db 11020 GGTCTCTCATGCT 11079  
Qy 1081 TCTCTCTCGAGGAGTATCTGAGGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140

Db 11080 TCTCTCTCGAGGAGTATCTGAGGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11139  
Qy 1141 GTCTCTCTGGGAGATTACCTTGTGCGGCCCAACAGACCGGGTGTCTACAGAACTC 1200  
Db 11140 GTCTCTCTGGGAGATTACCTTGTGCGGCCCAACAGACCGGGTGTCTACAGAACTC 11199  
Qy 1201 TGCAAGTTCCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCATCCCAAGGA 1260  
Db 11200 TGCAAGTTCCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCATCCCAAGGA 11259  
Qy 1261 CTGAGCACACCGGCATCCCACTGCTGTCAGAGGAGAGCCCTGACATCTCCTTTTTCAGACCT 1320  
Db 11260 CTGAGCACACCGGCATCCCACTGCTGTCAGAGGAGAGCCCTGACATCTCCTTTTTCAGACCT 11319  
Qy 1321 CATTCCTTCCACAGAGATGTTGAGAAATGTTTCCATCTCTCCAGCCCTGACCCCATGCTCTCT 1380  
Db 11320 CATTCCTTCCACAGAGATGTTGAGAAATGTTTCCATCTCTCCAGCCCTGACCCCATGCTCTCT 11379  
Qy 1381 GGAATCAGGCTCTGCTTCCCACTGTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG 1440  
Db 11380 GGAATCAGGCTCTGCTTCCCACTGTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGG 11439  
Qy 1441 AACATTTTCCAAAACCTGTCAGGGCGGGGTTGGTCTCAATCTCCTCGGGGCACTTTTCA 1500  
Db 11440 AACATTTTCCAAAACCTGTCAGGGCGGGGTTGGTCTCAATCTCCTCGGGGCACTTTTCA 11499  
Qy 1501 TCCTCAAGCTCAGGGCCCATCCCTCTCTGTCAGCTCTGACCCCAAAATTTAGTCCAGAAAT 1560  
Db 11500 TCCTCAAGCTCAGGGCCCATCCCTCTCTGTCAGCTCTGACCCCAAAATTTAGTCCAGAAAT 11559  
Qy 1561 AAATTCAGGAAG 1571  
Db 11560 AAATTCAGGAAG 11570

RESULT 5  
ACH74409  
ID ACH74409 standard; DNA; 567 BP.  
XX  
AC ACH74409;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
Human genome derived single exon probe #7604.  
XX  
Human; probe; ss; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
DR WPI; 2004-119264/12.  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
PS Claim 15; SEQ ID NO 7604; 80pp; English.  
XX

CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridises under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above). The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704  
CC

SQ Sequence 567 BP; 105 A; 197 C; 125 G; 140 T; 0 U; 0 Other;

Query Match 34.4%; Score 541; DB 12; Length 567;

Best Local Similarity 100.0%; Pred. No. 7e-128;

Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1031	GTCTCCTTCTGCCACATTGGCCATCTCTGCCTCTCATGCCCCCTTCTCTCTGTC	1090
Db	1	GTCTCCTTCTGCCACATTGGCCATCTCTGCCTCTCATGCCCCCTTCTCTCTGTC	60
Qy	1091	AGGCTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGTCTTGGG	1150
Db	61	AGGCTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGTCTTGGG	120
Qy	1151	GAGATTACCTTGTGCGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCA	1210
Db	121	GAGATTACCTTGTGCGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCA	180
Qy	1211	CCAGTGGATCCAGGAACCACTCAGGCCAATCTCTGAGTATCCAGGACTCAGCAC	1270
Db	181	CCAGTGGATCCAGGAACCACTCAGGCCAATCTCTGAGTATCCAGGACTCAGCAC	240
Qy	1271	CGGCATCCCACTGTCTCAGGGACAGCCCTGACACTCTTTTACAGCCCTCATTTCTTCC	1330
Db	241	CGGCATCCCACTGTCTCAGGGACAGCCCTGACACTCTTTTACAGCCCTCATTTCTTCC	300
Qy	1331	CAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTTACAGCCCTTCTCTGGACTCAGGG	1390
Db	301	CAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTTACAGCCCTTCTCTGGACTCAGGG	360
Qy	1391	TCTGTTCCCAACATTGGGCTGACCGGTCTCTCTAGTTGAAACCTTGGGAACAAATTTCC	1450
Db	361	TCTGTTCCCAACATTGGGCTGACCGGTCTCTCTAGTTGAAACCTTGGGAACAAATTTCC	420
Qy	1451	AAACTGTCCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTTCATCTTCAAGCT	1510
Db	421	AAACTGTCCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTTCATCTTCAAGCT	480

Qy	1511	CAGGGCCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGAA	1570
Db	481	CAGGGCCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGAA	540
Qy	1571	G 1571	
Db	541	G 541	
RESULT 6			
ABL63579/c			
ID	ABL63579	standard; DNA; 586 BP.	
XX	AC	ABL63579;	
XX	DT	15-MAY-2002 (first entry)	
XX	DE	Breast cancer related gene sequence SEQ ID NO:1916.	
XX	KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;	
XX	KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;	
XX	KW	cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;	
XX	OS	gene; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO200194629-A2.	
XX	PD	13-DEC-2001.	
XX	PF	30-MAY-2001; 2001WO-US010838.	
XX	PR	05-JUN-2000; 2000US-0209473P.	
XX	PR	05-JUN-2000; 2000US-0209531P.	
XX	PR	18-SEP-2000; 2000US-023133P.	
XX	PR	18-SEP-2000; 2000US-0231617P.	
XX	PR	20-SEP-2000; 2000US-0234009P.	
XX	PR	20-SEP-2000; 2000US-0234034P.	
XX	PR	20-SEP-2000; 2000US-0234052P.	
XX	PR	22-SEP-2000; 2000US-0234509P.	
XX	PR	22-SEP-2000; 2000US-0234567P.	
XX	PR	25-SEP-2000; 2000US-0234923P.	
XX	PR	25-SEP-2000; 2000US-0234924P.	
XX	PR	25-SEP-2000; 2000US-0235077P.	
XX	PR	25-SEP-2000; 2000US-0235082P.	
XX	PR	25-SEP-2000; 2000US-0235134P.	
XX	PR	25-SEP-2000; 2000US-0235280P.	
XX	PR	26-SEP-2000; 2000US-0235637P.	
XX	PR	26-SEP-2000; 2000US-0235638P.	
XX	PR	27-SEP-2000; 2000US-0235711P.	
XX	PR	27-SEP-2000; 2000US-0235720P.	
XX	PR	27-SEP-2000; 2000US-0235840P.	
XX	PR	27-SEP-2000; 2000US-0235863P.	
XX	PR	28-SEP-2000; 2000US-0236028P.	
XX	PR	28-SEP-2000; 2000US-0236032P.	
XX	PR	28-SEP-2000; 2000US-0236033P.	
XX	PR	28-SEP-2000; 2000US-0236034P.	
XX	PR	28-SEP-2000; 2000US-0236109P.	
XX	PR	28-SEP-2000; 2000US-0236111P.	
XX	PR	29-SEP-2000; 2000US-0236842P.	
XX	PR	29-SEP-2000; 2000US-0236891P.	
XX	PR	02-OCT-2000; 2000US-0237172P.	
XX	PR	02-OCT-2000; 2000US-0237173P.	
XX	PR	02-OCT-2000; 2000US-0237278P.	
XX	PR	02-OCT-2000; 2000US-0237294P.	
XX	PR	02-OCT-2000; 2000US-0237295P.	
XX	PR	02-OCT-2000; 2000US-0237316P.	
XX	PR	03-OCT-2000; 2000US-0237425P.	
XX	PR	03-OCT-2000; 2000US-0237598P.	
XX	PR	03-OCT-2000; 2000US-0237604P.	
XX	PR	03-OCT-2000; 2000US-0237606P.	
XX	PR	03-OCT-2000; 2000US-0237608P.	

PR 01-NOV-2000; 2000US-0244867P.  
XX 01-NOV-2000; 2000US-0245084P.  
PA (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
PS Claim 1; SEQ ID NO 1916; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
XX tumour  
XX  
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;  
  
Query Match 30.7%; Score 482; DB 6; Length 586;  
Best Local Similarity 100.0%; Pred. No. 8.3e-113;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1090 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGGGACTCGTGTCTCTGG 1149  
DB |||||  
QY 501 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGGGACTCGTGTCTCTGG 442  
DB |||||  
QY 1150 GGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGACCTCTGCAAGTTC 1209  
DB |||||  
QY 441 GGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGACCTCTGCAAGTTC 382  
DB |||||  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTTCATCCAGGACTCAGCACA 1269  
DB |||||  
QY 381 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTTCATCCAGGACTCAGCACA 322  
DB |||||  
QY 1270 CCGGCATCCCACTGTCTGAGGAGACGCGCTGACACTCTTTTCAGACCTTCATCTCTTC 1329  
DB |||||  
QY 321 CCGGCATCCCACTGTCTGAGGAGACGCGCTGACACTCTTTTCAGACCTTCATCTCTTC 262  
DB |||||  
QY 1330 CCAGAGATGTTGAGAAATTTTCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTCAGG 1389  
DB |||||  
QY 261 CCAGAGATGTTGAGAAATTTTCTCTCCAGCCCTGACCCCATGTCTCTGAGTCTCAGG 202  
DB |||||  
QY 1390 GTCTGTCTCCCACTGATGGGTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
DB |||||  
QY 201 GTCTGTCTCCCACTGATGGGTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 142  
DB |||||  
QY 1450 CAAAACCTGTCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGACTTTTCATCTCTCAAGC 1509  
DB |||||  
QY 141 CAAAACCTGTCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGACTTTTCATCTCTCAAGC 82  
DB |||||  
QY 1510 TCAGGGCCCATCTCTCTCTGACCTCTGACCCCAAAATTTAGTCCAGAAATAAACTGAGA 1569  
DB |||||  
QY 81 TCAGGGCCCATCTCTCTCTGACCTCTGACCCCAAAATTTAGTCCAGAAATAAACTGAGA 22  
DB |||||  
QY 1570 AG 1571

Db 21 AG 20

RESULT 7

ABL63996/c  
ID ABL63996 standard; DNA; 586 BP.  
XX AC ABL63996;  
XX  
DT 15-MAY-2002 (first entry)  
XX DE Breast cancer related gene sequence SEQ ID NO:2333.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0233133P.  
PR 18-SEP-2000; 2000US-0233617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 01-NOV-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
XX  
XX (AVAL-) AVALON PHARM.  
PA  
XX  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR., Weaver Z;  
XX WPI: 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
PT agent to be tested for anti-neoplastic activity, and determining a change  
PT in expression of a gene of a signature gene set.  
XX  
XX  
PS Claim 1; SEQ ID NO 2333; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an anti-  
CC neoplastic agent. The method involves exposing cells to a chemical agent  
CC to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening an  
CC anti-neoplastic agent, and can be used for producing a product which is  
CC the data collected with respect to the anti-neoplastic agent as a result  
CC of M1, and the data is sufficient to convey the chemical structure and/or  
CC properties of the agent. M1 can be used in the treatment of cancer such  
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms'  
XX tumour  
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;  
Query Match 30.7%; Score 482; DB 6; Length 586;  
Best Local Similarity 100.0%; Pred. No. 8.3e-113;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGAATTCGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149  
DB 501 CAGGGTGAATTCGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 442  
QY 1150 GAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209  
DB 441 GAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382  
QY 1210 ACCAAGTGGATCCAGGAACCTCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269  
DB 381 ACCAAGTGGATCCAGGAACCTCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 322  
QY 1270 CCGCATCCCCACCTGCTGAGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTTCTTC 1329  
DB 321 CCGCATCCCCACCTGCTGAGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTTCTTC 262  
QY 1330 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 1389  
DB 261 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 202  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 1449  
DB 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 142  
QY 1450 CAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTCATCTCTCAAGC 1509  
DB 141 CAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTCATCTCTCAAGC 82  
QY 1510 TCAGGGCCCATCCCTCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1569  
DB 81 TCAGGGCCCATCCCTCTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 22  
QY 1570 AG 1571  
DB 21 AG 20  
RESULT 8  
AAC79469/c

AAC79469 standard; cDNA; 735 BP.  
AAC79469;  
07-FEB-2001 (first entry)  
cDNA sequence of human breast tumour clone B541S.  
Human; breast tumour antigen; cytostatic; immunotherapy; breast cancer;  
vaccine; ss.  
Homo sapiens.  
W0200061756-A2.  
19-OCT-2000.  
10-APR-2000; 2000MO-US009688.  
09-APR-1999; 99US-00288950.  
02-JUL-1999; 99US-00346327.  
(CORI-) CORIXA CORP.  
Reed SG, Xu J, Dillon DC;  
WPI: 2000-638568/61.  
A novel isolated polypeptide comprising an immunogenic portion of a  
breast cancer protein useful in the detection and treatment of breast  
cancer.  
Claim 4; Page 88; 95pp; English.  
The present sequence was isolated from a breast tumour cDNA library. It  
is provided in a specification relating to compounds for immunotherapy  
and diagnosis of breast cancer. Breast tumour antigens and the  
polynucleotides that encode them may be used in the production of a  
pharmaceutical composition to be used in the treatment of breast cancer.  
Proliferated T cells and incubated antigen presenting cells are also  
required. The polypeptides and polynucleotides may also be used to  
produce a vaccine  
Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;  
Query Match 30.7%; Score 482; DB 3; Length 735;  
Best Local Similarity 100.0%; Pred. No. 9e-113;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGAATTCGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149  
DB 500 CAGGGTGAATTCGGGGGCGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 441  
QY 1150 GAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209  
DB 440 GAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 381  
QY 1210 ACCAAGTGGATCCAGGAACCTCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 1269  
DB 380 ACCAAGTGGATCCAGGAACCTCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGCACA 321  
QY 1270 CCGCATCCCCACCTGCTGAGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTTCTTC 1399  
DB 320 CCGCATCCCCACCTGCTGAGGAGACAGCCCTGACACTCCCTTTCAGACCCCTCATTTCTTC 261  
QY 1330 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 1389  
DB 260 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTGGACTCAGG 201  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 1449  
DB 200 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGAAACAATTC 141

QY 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509  
DB 140 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 81  
QY 1510 TCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCCAATTTTGTAGTCCAGAAATAAACTGAGA 1569  
DB 80 TCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCCAATTTTGTAGTCCAGAAATAAACTGAGA 21  
QY 1570 AG 1571  
DB 20 AG 19  
RESULT 9  
ID ABK29013/c  
XX 12-APR-2002 (first entry)  
AC ABK29013;  
XX Human breast tumour polypeptide cDNA clone #42.  
XX Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;  
KW immunostimulant.  
XX Homo sapiens.  
XX WO200198339-A2.  
XX 27-DEC-2001.  
XX 12-JUN-2001; 2001WO-US019032.  
XX 22-JUN-2000; 2000US-00602877.  
PR 12-OCT-2000; 2000US-00687507.  
PR 06-FEB-2001; 2001US-00778381.  
XX  
XX (CORI-) CORIXA CORP.  
XX Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;  
XX WPI; 2002-147792/19.  
XX Polynucleotides encoding breast tumor polypeptides, useful for treating  
PT breast cancer or stimulating an immune response.  
XX Claim 1; Page 140; 150pp; English.  
XX The invention relates to polynucleotides encoding breast tumour  
CC polypeptides. The sequences are useful for treating cancer, preferably  
CC breast cancer, in a patient or for stimulating an immune response. The  
CC polynucleotides and polypeptides are also useful in the diagnosis and  
CC monitoring of breast cancer. A method for detecting the presence of a  
CC cancer in a patient, comprises obtaining a biological sample from the  
CC patient, contacting the biological sample with a binding agent that binds  
CC to a breast tumour polypeptide, detecting in the sample an amount of  
CC polypeptide that binds to the binding agent, and comparing the amount of  
CC polypeptide to a predetermined cut-off value, therefore determining the  
CC presence of a cancer in the patient. Sequences ABK28920-ABK29025  
CC represent cDNA clones encoding human breast tumour polypeptides of the  
CC invention  
XX Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;  
SQ  
Query Match 30.7%; Score 482; DB 6; Length 735;  
Best Local Similarity 100.0%; Pred. No. 9e-113;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTATTCTGGGGGCGCTGCTGCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 1149  
DB 500 CAGGGTATTCTGGGGGCGCTGCTGCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 441

QY 1150 GGAGATTACCTTGTGTCGGGCCCAACAGACCGGGTGTCTACACGAACCTTCTCAAGTTC 1209  
DB 440 GGAGATTACCTTGTGTCGGGCCCAACAGACCGGGTGTCTACACGAACCTTCTCAAGTTC 381  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA 1269  
DB 380 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA 321  
QY 1270 CCGGCATCCCACTCTCTGAGGGCAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 1329  
DB 320 CCGGCATCCCACTCTCTGAGGGCAGCCCTGACACTCCCTTTCAGACCCCTCATTCCTTC 261  
QY 1330 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCGACTCAGG 1389  
DB 260 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTCGACTCAGG 201  
QY 1390 GTCTGCTTCCCAACATTTGGCTGACCGGTGTCTCTAGTTGAACCTTGGGAACAATTC 1449  
DB 200 GTCTGCTTCCCAACATTTGGCTGACCGGTGTCTCTAGTTGAACCTTGGGAACAATTC 141  
QY 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509  
DB 140 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 81  
QY 1510 TCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCCAATTTTGTAGTCCAGAAATAAACTGAGA 1569  
DB 80 TCAGGGCCCATCCCTTCTCTGCGAGCTCTGACCCCAATTTTGTAGTCCAGAAATAAACTGAGA 21  
QY 1570 AG 1571  
DB 20 AG 19  
RESULT 10  
AXX84240/c  
ID AAX84240 standard; cDNA; 738 BP.  
XX  
XX AAX84240;  
XX AC  
XX 08-SEP-1999 (first entry)  
XX DNA encoding human breast tumour protein immunogenic fragment.  
DE Breast tumour protein; immunogenic fragment; vaccine; detection;  
XX breast cancer development; therapy; ss.  
KW  
XX Homo sapiens.  
XX OS  
XX WO9933869-A2.  
XX 08-JUL-1999.  
XX 22-DEC-1998; 98WO-US027416.  
XX 24-DEC-1997; 97US-00998253.  
PR 24-DEC-1997; 97US-00998253.  
PR 17-JUL-1998; 98US-00118554.  
PR 17-JUL-1998; 98US-00118554.  
XX (CORI-) CORIXA CORP.  
XX Reed SG, Xu J;  
XX WPI; 1999-405486/34.  
XX New breast tumor protein genes used, in vaccines for immunotherapy, or  
PT for diagnosis of breast cancer.  
XX Claim 3; Page 70; 70pp; English.  
XX This sequence encodes a human breast tumour protein immunogenic fragment  
CC of the invention. The polypeptides or nucleic acids encoding them are  
CC useful in vaccines and pharmaceutical compositions for manufacture of

CC medicaments for inhibiting the development of breast cancer in a patient.  
CC They can also be used to treat breast cancer. Antibodies against these  
CC polypeptides can be used to detect and monitor progression of breast  
CC cancer in patients. Primers and probes derived from the polynucleotides  
CC encoding the breast proteins are useful for detection of breast cancer.  
CC Peripheral blood cells from a patient incubated in the presence of at  
CC least one polypeptide, such that T cells proliferate, are useful in  
CC manufacture of a medicament for treating breast cancer in a patient.  
CC Antigen presenting cells incubated in the presence of at least one  
CC polypeptide are also useful for treating breast cancer  
xx  
xx Sequence 738 BP; 161 A; 179 C; 226 G; 172 T; 0 U; 0 Other;

Query Match	30.7%	Score	482	DB 2	Length	738	
Best Local Similarity	100.0%	Prod. No.	9.1e-113				
Matches	482	Conservative	0	Mismatches	0	Gaps	0
Qy	1090	CAGGGTCATTCTGGGGGGCCCTGTGGTCTGCAAATGGCTCCCTGCGAGGACATCGTGTCTCTGG	1149				
Db	503	CAGGGTCATTCTGGGGGGCCCTGTGGTCTGCAAATGGCTCCCTGCGAGGACATCGTGTCTCTGG	444				
Qy	1150	GGAGATTACCTTGTGCGCCGCCCAAACAGACCCGGGTGTCTACACGAACCTCTGCAAGTTC	1209				
Db	443	GGAGATTACCTTGTGCGCCGCCCAAACAGACCCGGGTGTCTACACGAACCTCTGCAAGTTC	384				
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGGCAAATCTCTGAGTCACTCCAGGACTCAGCAC	1269				
Db	383	ACCAAGTGGATCCAGGAAACCATCCAGGGCAAATCTCTGAGTCACTCCAGGACTCAGCAC	324				
Qy	1270	CCGGCATCCCACTCTGTCAGGAGACGCCCTGCACATCCCTTTTCAGACCCCTCATTCCTTC	1329				
Db	323	CCGGCATCCCACTCTGTCAGGAGACGCCCTGCACATCCCTTTTCAGACCCCTCATTCCTTC	264				
Qy	1330	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCCCTGACCCCATGTCTCTGGGACTCAGG	1389				
Db	263	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCCCTGACCCCATGTCTCTGGGACTCAGG	204				
Qy	1390	GTCTGCTTCCCCACATGTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC	1449				
Db	203	GTCTGCTTCCCCACATGTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC	144				
Qy	1450	CAAAATGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGACATTCATCCTCAAGC	1509				
Db	143	CAAAATGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGACATTCATCCTCAAGC	84				
Qy	1510	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCGAAATAAATCAGAGA	1569				
Db	83	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCGAAATAAATCAGAGA	24				
Qy	1570	AG 1571					
Db	23	AG 22					

RESULT 11  
ADB80489  
ID ADB80489 standard; DNA; 1360 BP.  
XX  
XX  
XX AC ADB80489;  
AC  
XX  
XX 04-DEC-2003 (first entry)  
XX  
XX Ovarian cancer-associated transcript #27.  
XX  
XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;  
KW post-operative chemotherapy; radiation therapy; tumour prognosis;  
KW pre-cancerous lesion detection; ds; gene.

OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	43..924
FT		/*tag= a

XX	WO2002102235-A2.
XX	
XX	27-DEC-2002.
XX	
XX	18-JUN-2002; 2002WO-US019297.
XX	
XX	18-JUN-2001; 2001US-0299234P.
XX	18-JUN-2001; 2001US-0315287P.
XX	05-SEP-2001; 2001US-0317544P.
XX	13-NOV-2001; 2001US-0350666P.
XX	12-APR-2002; 2002US-0372246P.
XX	
XX	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
XX	Mack DH, Gish KC;
XX	
XX	WPI; 2003-167431/16.
XX	P-PSDB; ADB80490.
XX	
XX	Detecting an ovarian cancer-associated transcript in a cell from a
XX	patient, comprises contacting a biological sample from the patient with a
XX	polynucleotide that hybridizes to an ovarian cancer gene.
XX	
XX	Claim 10; Page 292; 332pp; English.
XX	
XX	The invention relates to a method of detecting an ovarian cancer-
XX	associated transcript in a cell from a patient, by contacting a
XX	biological sample from the patient with a polynucleotide that selectively
XX	hybridizes to a sequence at least 80% identical to any of one of 80
XX	nucleic acid sequences given in the specification. The method is useful
XX	in diagnosing ovarian cancer and in identifying and using agents and/or
XX	targets that inhibit ovarian cancer. The nucleic acid molecule,
XX	polypeptide and the antibody may also be used in detecting ovarian
XX	cancers, monitoring and early detection of relapse following treatment,
XX	monitoring response to therapy, selecting patients for post-operative
XX	chemotherapy or radiation therapy, in selecting mode of therapy,
XX	determining tumour prognosis, early detection of pre-cancerous lesions,
XX	and as vaccines. This sequence corresponds to one of the nucleic acids
XX	used for the detection method of the invention.

[illegible]



Db 1126 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCTCGGGGCACTTTTCATCTCTCAAGC 1185  
Qy 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1569  
Db 1186 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1245  
Qy 1570 AG 1571  
Db 1246 AG 1247  
RESULT 12  
ADN39197  
ID ADN39197 standard; cDNA; 1260 BP.  
XX  
AC ADN39197;  
XX  
DT 17-JUN-2004 (first entry)  
XX  
XX Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.  
XX  
XX Human; differential expression; cancer; angiogenic disorder;  
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
KW inflammatory disease; autoimmune disease;  
KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
KW detection; diagnosis; prognosis; drug screening; drug targeting;  
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
KW vulnery; gene therapy; vaccine; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2003042661-A2.  
XX  
XX 22-MAY-2003.  
XX  
XX 13-NOV-2002; 2002WO-US036810.  
XX  
XX 13-NOV-2001; 2001US-0350666P.  
PR 21-NOV-2001; 2001US-0332464P.  
PR 29-NOV-2001; 2001US-0334393P.  
PR 03-DEC-2001; 2001US-0335394P.  
PR 14-DEC-2001; 2001US-0340376P.  
PR 08-JAN-2002; 2002US-0347211P.  
PR 10-JAN-2002; 2002US-0347349P.  
PR 08-FEB-2002; 2002US-0352520P.  
PR 13-FEB-2002; 2002US-0356714P.  
PR 20-FEB-2002; 2002US-0359077P.  
PR 29-MAR-2002; 2002US-0368809P.  
PR 04-APR-2002; 2002US-0370110P.  
PR 12-APR-2002; 2002US-0372246P.  
PR 05-JUN-2002; 2002US-0386614P.  
PR 16-JUL-2002; 2002US-0396839P.  
PR 22-JUL-2002; 2002US-0397775P.  
PR 22-JUL-2002; 2002US-0397845P.  
PR 09-SEP-2002; 2002US-0409450P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
XX AFar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;  
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;  
XX  
XX WPI; 2003-468649/44.  
DR P-PSDB; ADN39198.  
XX  
XX Determining the presence or absence of a pathological cell in a patient,  
PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
PT a nucleic acid in a biological sample.  
XX  
XX Claim 8; SEQ ID NO 515; 1385pp; English.  
XX  
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
CC whose expression is upregulated or downregulated in specific cancers or  
CC other diseases such as angiogenic or fibrotic disorders, and to methods

CC of determining the presence or absence of a pathological cell in a  
CC patient by detecting a nucleic acid at least 80% identical to those of  
CC the invention or by detecting a polypeptide of the invention. The  
CC invention also relates to expression vectors and host cells comprising a  
CC nucleic acid of the invention; antibodies which specifically bind a  
CC polypeptide of the invention; use of such antibodies for drug targeting;  
CC and methods of screening for modulators of activity or expression of the  
CC polypeptides and nucleic acids. The nucleic acid, polypeptides,  
CC antibodies and methods are useful for diagnosing, prognosing and treating  
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,  
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
CC neovascularisation syndromes, scarring and uterine fibroids. They may  
CC also be useful in wound healing and in contraception. The present  
CC sequence represents a nucleic acid sequence of the invention.  
XX  
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;  
Query Match 30.7%; Score 482; DB 11; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 1.1e-112;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1090 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1149  
Db 766 CAGGGTGATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 825  
Qy 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209  
Db 826 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 885  
Qy 1210 ACCAATGGATCCAGGAACCATTCAGGGCCAACTCTCTAGTCTATCCAGGACTCAGACA 1269  
Db 886 ACCAATGGATCCAGGAACCATTCAGGGCCAACTCTCTAGTCTATCCAGGACTCAGACA 945  
Qy 1270 CCGGCATCCCACTCTGCGGGGAGAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329  
Db 946 CCGGCATCCCACTCTGCGGGGAGAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1005  
Qy 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1389  
Db 1006 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1065  
Qy 1390 GTCTGTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGAAACAATTC 1449  
Db 1066 GTCTGTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGAAACAATTC 1125  
Qy 1450 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509  
Db 1126 CAAAAGTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1185  
Qy 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1569  
Db 1186 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAAATAAACTGAGA 1245  
Qy 1570 AG 1571  
Db 1246 AG 1247  
RESULT 13  
ADR25550  
ID ADR25550 standard; DNA; 1260 BP.  
XX  
AC ADR25550;  
XX  
DT 21-OCT-2004 (first entry)  
XX  
XX Breast cancer prognosis marker #1411.  
DE ds; breast cancer; prognosis; gene expression; diagnosis.  
XX  
XX Homo sapiens.  
OS  
XX WO2004065545-A2.  
PN



XX 05-AUG-2004.  
XX PD  
XX PF 15-JAN-2004; 2004WO-US001100.  
XX PR 15-JAN-2003; 2003US-00342887.  
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.  
XX PA (NECA-) NETHERLANDS CANCER INST.  
XX PI Van't Veer LJ, He Y;  
XX DR WPI; 2004-593473/57.  
XX PT Classifying a breast cancer patient according to prognosis comprises  
PT determining the similarity between the level of expression of each of  
PT five genes in a cell sample taken from patient, to control levels.  
XX  
XX PS Disclosure; SEQ ID NO 1411; 226pp; English.  
XX  
CC The invention relates to a method of classifying a breast cancer patient  
CC according to prognosis by determining the similarity between the level of  
CC expression of each of five genes for which markers are listed in the  
CC specification, in a cell sample taken from the breast cancer patient, to  
CC control levels of expression for each respective five genes to obtain a  
CC patient similarity value. The methods are useful for classifying a breast  
CC cancer patient according to prognosis. Kits and computer program products  
CC are useful for data analysis using the diagnostic, prognostic and  
CC statistical methods of the invention. This sequence corresponds to a  
CC marker used in the method of the invention.  
XX  
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 13; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 1.1e-112;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGCTGATTCGGGGGGCTGTGCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 1149  
DB 766 CAGGCTGATTCGGGGGGCTGTGCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 825  
QY 1150 GGAGATTACCTCTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
DB 826 GGAGATTACCTCTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 885  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGACACA 1269  
DB 886 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGACACA 945  
QY 1270 CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCTTTTTCAGACCCCTCATCTCTTC 1329  
DB 946 CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCTTTTTCAGACCCCTCATCTCTTC 1005  
QY 1330 CCAGAGATTGTGAGAAATTCATCTCTCCAGCCCTGACCCCATGTCTCTGGAATCAGG 1389  
DB 1006 CCAGAGATTGTGAGAAATTCATCTCTCCAGCCCTGACCCCATGTCTCTGGAATCAGG 1065  
QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 1449  
DB 1066 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 1125  
QY 1450 CAAAACCTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1509  
DB 1126 CAAAACCTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1185  
QY 1510 TCAGGGCCATCCCTTCTCTGACCTGTGACCCCAAAATTTAGTCCAGAAATAAATCAGAGA 1569  
DB 1186 TCAGGGCCATCCCTTCTCTGACCTGTGACCCCAAAATTTAGTCCAGAAATAAATCAGAGA 1245  
QY 1570 AG 1571  
DB 1246 AG 1247

## RESULT 14

AA57990  
ID AAX57990 standard; DNA; 1302 BP.  
XX  
XX AC AAX57990;  
XX  
XX DT 19-JUL-1999 (first entry)  
XX DE Human BS247 specific polynucleotide #8.  
XX KW BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
XX KW fibroadenoma; cystic breast disease; gene therapy; ss.  
XX OS Homo sapiens.  
XX PN WO9922027-A1.  
XX PD 06-MAY-1999.  
XX PF 28-OCT-1998; 98WO-US022906.  
XX PR 28-OCT-1997; 97US-0063431P.  
XX PR 28-OCT-1997; 97US-00968838.  
XX PA (ABBO ) ABBOTT LAB.  
XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;  
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;  
PI Stroupe SD, Yu H;  
XX WPI; 1999-312977/26.  
XX  
XX Breast tissue derived cDNA contig and consensus polypeptide sequence.  
XX Claim 1; Page 106; 112pp; English.  
XX  
CC This sequence is a BS247 specific polynucleotide. The invention relates  
CC to a method of detecting the presence of a target BS247 polynucleotide,  
CC especially mRNA, in a test sample. BS247 polynucleotides are derived from  
CC breast tissue. The polynucleotides, polypeptides or antibodies are useful  
CC for providing information leading to the detection, diagnosis, staging,  
CC monitoring, prognosis, in vivo imaging, prevention or treatment,  
CC determining predisposition to, diseases and conditions of the breast,  
CC such as breast cancer, atypical hyperplasia, fibroadenoma and cystic  
CC breast disease. Drug treatment or gene therapy for breast cancer, can be  
CC based on these identified gene sequences and the efficacy of any  
CC particular therapy can be monitored. The BS247-derived reagents are  
CC advantageous for detection of breast cancer due to their specificity. The  
CC reagents also provide an alternative, non-surgical diagnostic method  
CC capable of detecting early stage breast disease, such as cancer  
XX  
SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;  
Query Match 30.7%; Score 482; DB 2; Length 1302;  
Best Local Similarity 100.0%; Pred. No. 1.1e-112;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGCTGATTCGGGGGGCTGTGCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 1149  
DB 811 CAGGCTGATTCGGGGGGCTGTGCTGCAATGGCTCCCTGAGGAGCTCGTCTCTGG 870  
QY 1150 GGAGATTACCTCTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
DB 871 GGAGATTACCTCTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 930  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGACACA 1269  
DB 931 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGACACA 990  
QY 1270 CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCTTTTTCAGACCCCTCATCTCTTC 1329  
DB 991 CCGGCATCCCACTGCTGAGGAGACGCCCTGACACTCTTTTTCAGACCCCTCATCTCTTC 1050

This sequence is a BS247 specific polynucleotide. The invention relates to a method of detecting the presence of a target BS247 polynucleotide, especially mRNA, in a test sample. BS247 polynucleotides are derived from breast tissue. The polynucleotides, polypeptides or antibodies are useful for providing information leading to the detection, diagnosis, staging, monitoring, prognosis, in vivo imaging, prevention or treatment, determining predisposition to, diseases and conditions of the breast, such as breast cancer, atypical hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or gene therapy for breast cancer, can be based on these identified gene sequences and the efficacy of any particular therapy can be monitored. The BS247-derived reagents are advantageous for detection of breast cancer due to their specificity. The reagents also provide an alternative, non-surgical diagnostic method capable of detecting early stage breast disease, such as cancer

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 178.302 Seconds

(without alignments)  
14417.061 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_10000\_11570

Perfect score: 1571

Sequence: 1 agggagtgagggaattga.....cccgagaataaactgagaag 1571

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1567.8	99.8	10818	4	US-09-949-016-13583
2	599	38.1	601	4	US-09-949-016-63416
3	482	30.7	735	3	US-09-602-877A-94
4	482	30.7	1143	4	US-09-949-016-1841
5	482	30.7	1499	4	US-09-509-908-1
6	480.6	30.6	601	4	US-09-949-016-63415
7	476	30.3	1476	2	US-08-824-874-2
8	476	30.3	1476	3	US-09-210-084-2
9	476	30.3	1476	4	US-09-764-762-2
10	460	29.3	1504	3	US-09-280-116-1
11	253	16.1	601	4	US-09-949-016-63414
12	223.4	14.2	68719	4	US-09-949-016-12799
13	223.4	14.2	68720	4	US-09-949-016-14296
14	222.8	14.2	32278	4	US-09-949-016-14575
15	222.6	14.2	176373	3	US-09-128-155-17
16	222.2	14.1	601	4	US-09-949-016-66914
17	222.2	14.1	177669	4	US-09-949-016-13713
18	221.8	14.1	80858	4	US-09-949-016-12659
19	221.8	14.1	80859	4	US-09-949-016-15715
20	220.8	14.1	69701	4	US-09-949-016-14187
21	220.8	14.1	73308	4	US-09-949-016-16326
22	219.8	14.0	300402	4	US-09-949-016-13632
23	219.4	14.0	38009	4	US-09-949-016-13617
24	219.4	14.0	70000	3	US-09-851-896-3
25	219.4	14.0	76399	4	US-09-949-016-16819
26	219.2	14.0	152331	3	US-09-128-155-16
27	218.8	13.9	71574	4	US-09-949-016-15580

Sequence 4, Appli  
Sequence 15691, A  
Sequence 15728, A  
Sequence 13664, A  
Sequence 13969, A  
Sequence 12128, A  
Sequence 13261, A  
Sequence 13466, A  
Sequence 13113, A  
Sequence 11852, A  
Sequence 14166, A  
Sequence 14293, A  
Sequence 16474, A  
Sequence 14902, A  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 1, Appli

US-09-326-480A-4  
US-09-949-016-15691  
US-09-949-016-15728  
US-09-949-016-13664  
US-09-949-016-13969  
US-09-949-016-12128  
US-09-949-016-13261  
US-09-949-016-13466  
US-09-949-016-13113  
US-09-949-016-11852  
US-09-949-016-14166  
US-09-949-016-14293  
US-09-949-016-16474  
US-09-949-016-14902  
US-09-404-879A-1  
US-09-404-879A-3  
US-09-338-933-1

## ALIGNMENTS

## RESULT 1

US-09-949-016-13583  
Sequence 13583, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: Fast-Seq for Windows Version 4.0

SEQ ID NO 13583

LENGTH: 10818

TYPE: DNA

ORGANISM: Human

US-09-949-016-13583

Query Match 99.8%; Score 1567.8; DB 4; Length 10818;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGGGAGTATGGGAATTGGAAGACAGGAAACACAAATAGTCCAAAGCAATGGATTCTA 60  
DB 7243 AGGGAGTATGGGAATTGGAAGACAGGAAACACAAATAGTCCAAAGCAATGGATTCTA 7302

QY 61 TTGGAGTATGTTGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGTGTTCA 120  
DB 7303 TTGGAGTATGTTGCCCCCTAGGAAGACACTGGCAATACCAAGGAGACATTTTGTGTTCA 7362

QY 121 CAATATATGGAGGGGCAATCTACTGGCACTAATGATAGATGCAAGTGTCTGTTCAAC 180  
DB 7363 CAATATATGGAGGGGCAATCTACTGGCACTAATGATAGATGCAAGTGTCTGTTCAAC 7422

QY 181 ATGCTATGATGCACGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC 240  
DB 7423 ATGCTATGATGCACGCGAGGCTCCACAAACCAATTCAGCTTCAGATGCCAC 7482

QY 241 AGTCCAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG 300  
DB 7483 AGTCCAGATCGAGGAACCCCTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG 7542

QY 301 GTATAAGGATGGGTTGGTGGAGAAATGGGAAGGAGTGTGTGTCAGTAAGAAATAA 360

Db 7543 GTATAAGGATGGGTGGTGGAGAAATGGGGAAGGAGGTGCTGTCTCCAGTAGTAAGAAATAA 7602  
QY  
Db 361 GGCCTGCACAGGCTGGAGGGAGAGTCAGAGAGAAAGGGAGGGGAGAGATACACGATGA 420  
Db 7603 GGCCTGCACAGGCTGGAGGGAGAGTCAGAGAGAAAGGGAGGGGAGAGATACACGATGA 7662  
QY 421 GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAAGATTCGAGATGCTGGAGAGGAAGGGTCA 480  
Db 7663 GGGAGACAGGCTGGAAACAGAAAGTAGAGACGAAGATTCGAGATGCTGGAGAGGAAGGGTCA 7722  
QY 481 CAGACCCCCCGAAATGATGTGGCAACACAGAAATCTGGAAAGGAAGATGGAGTGGAG 540  
Db 7723 CAGACCCCCCGAAATGATGTGGCAACACAGAAATCTGGAAAGGAAGATGGAGTGGAG 7782  
QY 541 AGTGACAAATGGGGTCTAAAGGTTGAATTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 600  
Db 7783 AGTGACAAATGGGGTCTAAAGGTTGAATTTGGAGGCCAGGCATGGTGGCTCAGCCCTGTA 7842  
QY 601 ATCCCAACACTTTGGAGGCTGAGGTGGGCGAAATCACTTGAGGCCAGGAGTTCGAGACCAG 660  
Db 7843 ATCCCAACACTTTGGAGGCTGAGGTGGGCGAAATCACTTGAGGCCAGGAGTTCGAGACCAG 7902  
QY 661 CTTGGCCAAATGGTGAAACCCCGTCTCTACAAAAAATACAAAAAATTTAGCCGGGTGT 720  
Db 7903 CTTGGCCAAATGGTGAAACCCCGTCTCTACAAAAAATACAAAAAATTTAGCCGGGTGT 7962  
QY 721 GGTGTGAGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCAGGAGAAATTCCTTGAAC 780  
Db 7963 GGTGTGAGACACCTGTAGTCACAGCTACTTTGGAGGCTGAGGCAGGAGAAATTCCTTGAAC 8022  
QY 781 CCGGGAGATGGAGCTGCAGTGAGCTGAGGTCAGGCCACTGCGCTCCAACTCGGGCAACA 840  
Db 8023 CTGGAGATGGAGCTGCAGTGAGCTGAGGTCAGGCCACTGCGCTCCAACTCGGGCAACA 8082  
QY 841 GAGTAAGACTCCATCTCAAAAAAATAAAGCTGGAATTTGAGGTGAATATTAATAACATT 900  
Db 8083 GAGTAAGACTCCATCTCAAAAAAATAAAGCTGGAATTTGAGGTGAATATTAATAACATT 8142  
QY 901 CTCCTCTCTCTCCTTTTGCTGCTGCTCCATCTGCTCTTTTTCGCAATTTCTCATCT 960  
Db 8143 CTCCTCTCTCTCCTTTTGCTGCTGCTCCATCTGCTCTTTTTCGCAATTTCTCATCT 8202  
QY 961 CTGTACTTTCCATCTCTGTGTCTGTCTCCATCTGTCTCTCCATCTATGGGCATCTCTG 1020  
Db 8203 CTGTACTTTCCATCTCTGTGTCTGTCTCCATCTGTCTCTCATCTATGGGCATCTCTG 8262  
QY 1021 GGTCTCTCATGTCTCTTCTGGCCACATTTGGCCACATCTCTGCTCTCTCATGCCCCCTT 1080  
Db 8263 GGTCTCTCATGTCTCTTCTGGCCACATTTGGCCACATCTCTGCTCTCTCATGCCCCCTT 8322  
QY 1081 TCTCTCTCGCAGGCTGATTTCTGGGGGCTGTGTGTCTGCAATGGCTCCCTGCAAGGACTC 1140  
Db 8323 TCTCTCTCGCAGGCTGATTTCTGGGGGCTGTGTGTCTGCAATGGCTCCCTGCAAGGACTC 8382  
QY 1141 GTGTCTCTGGGAGATTAACCTTTGTCGGGCCCCAACAGACCGGGTGTCTACAGAAACCTC 1200  
Db 8383 GTGTCTCTGGGAGATTAACCTTTGTCGGGCCCCAACAGACCGGGTGTCTACAGAAACCTC 8442  
QY 1201 TGCAAGTTTCAACAAAGTGATTCAGGAAACCAATCAGGCCAACTCCTGAGTCATCCAGGA 1260  
Db 8443 TGCAAGTTTCAACAAAGTGATTCAGGAAACCAATCAGGCCAACTCCTGAGTCATCCAGGA 8502  
QY 1261 CTGAGCACACCGCATCCCACTGTCTGAGGAGACGCTGACACTCTTTCAGACCTT 1320  
Db 8503 CTGAGCACACCGCATCCCACTGTCTGAGGAGACGCTGACACTCTTTCAGACCTT 8562  
QY 1321 CATTCCTTCCAGAGATGTTGAGAATGTTTCACTCTCCAGCCCCCTGACCCCATGCTCTCT 1380  
Db 8563 CATTCCTTCCAGAGATGTTGAGAATGTTTCACTCTCTCCAGCCCCCTGACCCCATGCTCTCT 8622  
QY 1381 GGAATCAGGGTCTGCTTCCCCCAATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTTGGG 1440

Db 8623 GGAATCAGGGTCTGCTTCCCCCAATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTTGGG 8682  
QY 1441 AACAAATTTCCAAACTGTCTCAGGGCGGGGTTCGCTCTCAATCTCCTGGGGGACTTTCA 1500  
Db 8683 AACAAATTTCCAAACTGTCTCAGGGCGGGGTTCGCTCTCAATCTCCTGGGGGACTTTCA 8742  
QY 1501 TCCTCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCTGACCCAAATTTAGTCCAGAAAT 1560  
Db 8743 TCCTCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCTGACCCAAATTTAGTCCAGAAAT 8802  
QY 1561 AAATCAGGAAG 1571  
Db 8803 AAATCAGGAAG 8813

RESULT 2  
US-09-949-016-63416/c  
; Sequence 63416, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63416  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63416

Query Match 38.1%; Score 599; DB 4; Length 601;  
Best Local Similarity 99.7%; Pred. No. 3.4e-151;  
Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 482 AGACCCCCCGGAAATGATGTGGCAACACAGGAATCTGGAGAGGAGATGGAGTGGAGA 541  
Db 601 AGACCCCCCGGAAATGATGTGGCAACACAGGAATCTGGAGAGGAGATGGAGTGGAGA 542  
QY 542 GTGACAAATGGGCTCTAAAGGTTGAATTTGGAGGCCAGGCATGTGGCTCACGCCCTGTA 601  
Db 541 GTGACAAATGGGCTCTAAAGGTTGAATTTGGAGGCCAGGCATGTGGCTCACGCCCTGTA 482  
QY 602 TCCAAACATTTGGAGGCTGAGTGGGCGAATCACTTGAGGCCAGGAGTTCGAGACACAG 661  
Db 481 TCCAAACATTTGGAGGCTGAGTGGGCGAATCACTTGAGGCCAGGAGTTCGAGACACAG 422  
QY 662 CTGGCCCAACATGCTGAAACCCCGTCTCTACAAAAAATAACAAAAATTTAGCCGGGTGTG 721  
Db 421 CTGGCCCAACATGCTGAAACCCCGTCTCTACAAAAAATAACAAAAATTTAGCCGGGTGTG 362  
QY 722 GTGATGGACACCTGTAGTCACTGCTTGGAGGCTGAGGCCAGGAGAAATTTGCTTGAACC 781  
Db 361 GTGATGGACACCTGTAGTCACTGCTTGGAGGCTGAGGCCAGGAGAAATTTGCTTGAACC 302  
QY 782 CGGGAGATGGAGGCTGCTGAGTCACTGAGGCTCAGGCCAATGCGCTCCAACTCGGGCAACAG 841  
Db 301 YGGGAGATGGAGGCTGCTGAGTCACTGAGGCTCAGGCCAATGCGCTCCAACTCGGGCAACAG 242  
QY 842 AGTAAGACTCCATCTCAAAAAAATAAAGCTGATTTGGAGTCAAAATATTAATAACATTC 901  
Db 241 AGTAAGACTCCATCTCAAAAAAATAAAGCTGATTTGGAGTCAAAATATTAATAACATTC 182  
QY 902 TCCCTCTCTCTCTTTTGGCTGTGCTCCATCTCTCTCTTTTTCGCAATTTCTTCATCTC 961

Db	181	TCCCTCTCTCTCCCTTTGGCGTGCTGCCATCTCTGCTTTTCTCGATTCTTCACTC	122
Qy	962	TGTACTTTCCATCTCTGTGTCTGTTCCCATCTGTCTTCCATCTATGGCATCTCTGG	1021
Db	121	TGTACTTTCCATCTCTGTGTCTGTTCCCATCTGTCTTCCATCTATGGCATCTCTGG	62
Qy	1022	GTCTCTCATGTCTCCTTCTGCCCACTTTGCCACACTCTCTGCGCTCTCTCATGCCCCCCTTT	1081
Db	61	GTCTCTCATGTCTCCTTCTGCCCACTTTGCCACACTCTCTGCGCTCTCTCTGCCCCCCTTT	2
Qy	1082	C 1082	
Db	1	C 1	

```

RESULT 3
US-09-602-877A-94/c
; Sequence 94, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 94
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-94

```

Query Match	30.7%;	Score 482;	DB 3;	Length 735;
Best Local Similarity	100.0%;	Prod. No. 1.1e-119;		
Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1090	CAGGGTGATTCTGGGGGCGCTGTGGTCTGCNAATGGCTCCCTGCAGGGACTCGTGTCTCTGG	1149	
Db	500	CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCTGG	441	
Qy	1150	GGAGATTACCCCTTGTGCGCGGCCCAAAGACACGGGGTGCTACACGAACTCTCGAAGTTC	1209	
Db	440	GGAGATTACCCCTTGTGCGCGGCCCAAAGACACGGGGTGCTACACGAACTCTCGAAGTTC	381	
Qy	1210	ACCAAGTGGATCCAGGAAACCAATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCAC	1259	
Db	380	ACCAAGTGGATCCAGGAAACCAATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCAC	321	
Qy	1270	CCGGCATCCCAACTGTGTGAGGACAGCCCTGACATCTCTTTTCAGACCTCATTCCTTC	1329	
Db	320	CCGGCATCCCAACTGTGTGAGGACAGCCCTGACATCTCTTTTCAGACCTCATTCCTTC	261	
Qy	1330	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCGTACCCCATGTCTCTGGACTCAGG	1389	
Db	260	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCGTACCCCATGTCTCTGGACTCAGG	201	
Qy	1390	GTCTGCTTTCCCCACATTGGGCTCAGCGTGTCTCTAGTTTGAACCCCTGGGAAACAAATTC	1449	
Db	200	GTCTGCTTTCCCCACATTGGGCTCAGCGTGTCTCTAGTTTGAACCCCTGGGAAACAAATTC	141	
Qy	1450	CAAAACATGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1509	
Db	140	CAAAACATGTCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	81	
Qy	1510	TCAGGGCCCAATCCCTTCTCTGCACTGTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569	
Db	80	TCAGGGCCCAATCCCTTCTCTGCACTGTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	21	

Cy	1570 AG 1571	
Db	20 AG 19	

RESULT 4  
 US-09-949-016-1841  
 ; Sequence 1841, Application US/09949016  
 ; Patent NO. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CLO01307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1841  
 ; LENGTH: 1143  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-1841

Query Match	30.7%	Score 482	DB 4	Length 1143	
Best Local Similarity	100.0%	Prod. No. 1.3e-119			
Matches	482	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1090	CAGGGTGATTTCTGGGGGCCCTTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGCTCTGG	1149		
Db	657	CAGGGTGATTTCTGGGGGCCCTTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGCTCTGG	716		
Qy	1150	GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTTC	1209		
Db	717	GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTTC	776		
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGACACA	1269		
Db	777	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGACACA	836		
Qy	1270	CCGGCATCCCCACCTGCTGCAGGGAGACGCCCTGCACACTCCTTTACAGCCCTCATTCCTTC	1329		
Db	837	CCGGCATCCCCACCTGCTGCAGGGAGACGCCCTGCACACTCCTTTACAGCCCTCATTCCTTC	896		
Qy	1330	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCCCTGACCCCATGTCTCTCGGACTCAG	1389		
Db	897	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCCCTGACCCCATGTCTCTCGGACTCAG	956		
Qy	1390	GTCTGCTTCCCCACATATGGGCTGACCGGTCTCTCTAGTTGAAACCTGGGAAACAAATTC	1449		
Db	957	GTCTGCTTCCCCACATATGGGCTGACCGGTCTCTCTAGTTGAAACCTGGGAAACAAATTC	1016		
Qy	1450	CAAACTGTCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1509		
Db	1017	CAAACTGTCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1076		
Qy	1510	TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569		
Db	1077	TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1136		
Qy	1570	AG 1571			
Db	1137	AG 1138			

RESULT 5  
US-09-509-908-1

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
00

1 Application US/09509908  
2  
3 GENERAL INFORMATION:  
4  
5 APPLICANT: The Procter & Gamble Company, N/A N/A  
6  
7 TITLE OF INVENTION: A Protease  
8  
9 NUMBER OF SEQUENCES: 2  
10  
11 CORRESPONDENCE ADDRESS:  
12  
13 ADDRESSEE: T. David Reed  
14  
15 STREET: 5299 Spring Grove Avenue  
16  
17 CITY: Cincinnati  
18  
19 STATE: Ohio  
20  
21 COUNTRY: USA  
22  
23 ZIP: 45217-1087  
24  
25 COMPUTER READABLE FORM:  
26  
27 MEDIUM TYPE: Floppy disk  
28  
29 OPERATING: IBM PC compatible  
30  
31 OPERATING SYSTEM: PC-DOS/MS-DOS  
32  
33 SOFTWARE: PatentIn Release #1.0, Version #1.30  
34  
35 CURRENT APPLICATION DATA:  
36  
37 APPLICATION NUMBER: US/09/509,908  
38  
39 FILING DATE: 28-Feb-2002  
40  
41 CLASSIFICATION: <Unknown>  
42  
43 ATTORNEY/AGENT INFORMATION:  
44  
45 NAME: Reed, T David  
46  
47 REGISTRATION NUMBER: 32,931  
48  
49 REFERENCE/DOCKET NUMBER: AA-264F  
50  
51 TELECOMMUNICATION INFORMATION:  
52  
53 TELEPHONE: 513-627-7025  
54  
55 TELEFAX: 513-627-6333

```

/ INFORMATION FOR SEQ ID NO: 1:
/   SEQUENCE CHARACTERISTICS:
/       LENGTH: 1499 base pairs
/       TYPE: nucleic acid
/       STRANDEDNESS: single
/       TOPOLOGY: linear
/   MOLECULE TYPE: DNA (genomic)
/   FEATURE:
/       NAME/KEY: CDS
/       LOCATION: 291..1172
/   FEATURE:
/       NAME/KEY: mat_peptide
/       LOCATION: 489..1172
/   FEATURE:
/       NAME/KEY: 5'UTR
/       LOCATION: 1..290
/   FEATURE:
/       NAME/KEY: 3'UTR
/       LOCATION: 1173..1499
/   SEQUENCE DESCRIPTION: SEQ ID
/
US-09-509-908-1

```

Query Match	30.7%	Score 482;	DB 4;	Length 1499;
Best Local Similarity	100.0%;	Prod. No. 1.5e-119;		
Matches 482; Conservative 0;	Mismatches 119;	Indels 0;	Gaps 0;	
QY	1090	CAGGGTGATTCTCGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	1149	
Db	1014	CAGGGTGATTCTCGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	1073	
QY	1150	GGAGATTACCCCTTGTGCGCGGCCCAACAGACCCGGGTGCTTACAGCAACCTCTGCAAGTTC	1209	
Db	1074	GGAGATTACCCCTTGTGCGCGGCCCAACAGACCCGGGTGCTTACAGCAACCTCTGCAAGTTC	1133	
QY	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTGAGTGCATCCAGGACTCAGCAC	1369	
Db	1134	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTGAGTGCATCCAGGACTCAGCAC	1193	
QY	1270	CCGGCATCCCCACTGCTGCAGGAGACGCCCTGCACACTCCCTTTCAGACCCCTCATTCCTTC	1329	
Db	1194	CCGGCATCCCCACTGCTGCAGGAGACGCCCTGCACACTCCCTTTCAGACCCCTCATTCCTTC	1353	
QY	1330	CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCCCTGACCCCACTGTCTCTGGACTCAGG	1389	

Db	1254	CCAGAGATGTTGAGAAATGTTCACTCTCTCAGCCCTGACCCCAATGTCCTCTGGAGCTCAGG	1313
Qy	1390	GTCTGCTTCCCCACAAATTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAAATTTTC	1449
Db	1314	GTCTGCTTCCCCACAAATTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAAATTTTC	1373
Qy	1450	CAAAATGTCACGGGCGGGGTTGCGTCTCAATCTCCCTGGGGGCACTTTTCATCCTCAAGC	1509
Db	1374	CAAAATGTCACGGGCGGGGTTGCGTCTCAATCTCCCTGGGGGCACTTTTCATCCTCAAGC	1433
Qy	1510	TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCGAGAAATAAACTGAGA	1569
Db	1434	TCAGGGCCCATCCCTTCTCTGACGCTGTGACCCAAATTTAGTCCGAGAAATAAACTGAGA	1493
Qy	1570	AG 1571	
Db	1494	AG 1495	

## RESULT 6

```

US-09-949-016-63415/C
; Sequence 63415, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fasteq for Windows Version 4.0
; SEQ ID NO 63415
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63415

```

Query Match	30.6%	Score	480.6	DB	4	Length	601
Best Local Similarity	99.8%	Pred. No.	2.3e-119				
Matches	480	Conservative	1	Mismatches	0	Indels	0
Gaps	0						
Qy	1	AGGGAGTGATGGGGAATTGAAGACAGGAACACAAATTTAGTCCAAAGCGAATGGATTTCTA	60				
Db	481	AGGGAGTGATGGGGAATTGAAGACAGGAACACAAATTTAGTCCAAAGCGAATGGATTTCTA	422				
Qy	61	TTGGGAGTGATTTCTGCCCTTAGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	120				
Db	421	TTGGGAGTGATTTCTGCCCTTAGAAGACACTGGCAATACCAAGGAGACATTTTGGTTGTCA	362				
Qy	121	CAACTATATGGAGGGGCAATTACTTGGCAATPAATGGATAGATGCCAAGTGTCTGTTCAAC	180				
Db	361	CAACTATATGGAGGGGCAATTACTTGGCAATPAATGGATAGATGCCAAGTGTCTGTTCAAC	302				
Qy	181	ATGCTATGATGCACAGGACGGCTCCACAAACAAACCATTTATCCAGCTTCAGATGCCAC	240				
Db	301	ATGCTATGATGCACAGGACGGCTCCACAAACAAACCATTTATCCAGCTTCAGATGCCAC	242				
Qy	241	AGTGCCCGATCAGGAAACCCCTCATCAGGGGCTGAGAACCGTATTTTTCAGAAAGGGAG	300				
Db	241	AGTGCCCGATCAGGAAACCCCTCATCAGGGGCTGAGAACCGTATTTTTCAGAAAGGGAG	182				
Qy	301	GTATAGGATGGGTTGGTGGAGAAATGGGGAAGAAAGGTGTGTCTCCAGTAAAGAGAAATAA	360				
Db	181	GTATAGGATGGGTTGGTGGAGAAATGGGGAAGAAAGGTGTGTCTCCAGTAAAGAGAAATAA	122				
Qy	361	GGCCTGCACAGGCTGGAGGGGAGAGTTCAGAGAGAAAGGGAGCGGAGAGATACACGATGA	420				

Db 121 GGCCTGCACAGGCTGGAGGGGAGAGTGCAGAGAGAAAGGAGCGGAGAGATACACAGTGA 62  
Qy 421 GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTCGAGATGTGGAGAGGAAGGTCA 480  
Db 61 GGGAGACAGGCTGGAACAGAAAGTAGAGACGAAGATTCGAGATGTGGAGAGGAAGGTCA 2  
Qy 481 C 481  
Db 1 C 1

## RESULT 7

US-08-874-2  
; Sequence 2, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA

; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1476 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: KERANOT02  
; CLONE: 820694

US-08-874-2  
Query Match 30.3%; Score 476; DB 2; Length 1476;  
Best Local Similarity 100.0%; Pred. No. 6e-118;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1090 CAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTGG 1149  
Db 1001 CAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTGG 1060  
Qy 1150 GGAGATTACCTTGTGCCCCCGGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTC 1209  
Db 1061 GGAGATTACCTTGTGCCCCCGGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTC 1120  
Qy 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGAGACTCAGACACA 1269  
Db 1121 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGAGACTCAGACACA 1180

Qy 1270 CCGGCATCCCACTGCTGCAGGGAGAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1329  
Db 1181 CCGGCATCCCACTGCTGCAGGGAGAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240  
Qy 1330 CCAGAGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1389  
Db 1241 CCAGAGATGTTGAGAAATGTTCACTCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1300  
Qy 1390 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 1449  
Db 1301 GTCTGCTTCCCACTGAGTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1360  
Qy 1450 CAAACTGTCCAGGGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC 1509  
Db 1361 CAAACTGTCCAGGGGGGGTTCGGTCTCAATCTCCCTGGGCACTTTTCATCTCTCAAGC 1420  
Qy 1510 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACT 1565  
Db 1421 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACT 1476

## RESULT 8

US-09-210-084-2  
; Sequence 2, Application US/09210084  
; Patent No. 6197511  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,084  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1476 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: KERANOT02  
; CLONE: 820694

US-09-210-084-2  
Query Match 30.3%; Score 476; DB 3; Length 1476;  
Best Local Similarity 100.0%; Pred. No. 6e-118;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1090 CAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTGG 1149



Db 1001 ||||| CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTCTGG 1060  
QY 1150 GGAGATTACCTTGTGTCGGGCCCCAACAGACGGGTGCTTACACGAACTCTCGAAGTTC 1209  
Db 1061 GGAGATTACCTTGTGTCGGGCCCCAACAGACGGGTGCTTACACGAACTCTCGAAGTTC 1120  
QY 1210 ACCAAGTGAATCCAGGAAACATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1269  
Db 1121 ACCAAGTGAATCCAGGAAACATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1180  
QY 1270 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTC 1329  
Db 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTC 1240  
QY 1330 CCAGAGATTGAGATGTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGACTCAGG 1389  
Db 1241 CCAGAGATTGAGATGTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGACTCAGG 1300  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
Db 1301 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1360  
QY 1450 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCCCTGGGACACTTTTCATCTCAAGC 1509  
Db 1361 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCCCTGGGACACTTTTCATCTCAAGC 1420  
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACT 1565  
Db 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACT 1476

RESULT 9

US-09-764-762-2  
; Sequence 2, Application US/09764762  
; Patent No. 6472195  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/764,762  
; FILING DATE: 16-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/210,084  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1476 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; IMMEDIATE SOURCE:  
; LIBRARY: KERANOT02  
; CLONE: 820694  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-764-762-2  
  
Query Match 30.3%; Score 476; DB 4; Length 1476;  
Best Local Similarity 100.0%; Pred. No. 6e-118;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1090 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1149  
Db 1001 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1060  
QY 1150 GGAGATTACCTTGTGTCGGGCCCCAACAGACGGGTGCTTACACGAACTCTCGAAGTTC 1209  
Db 1061 GGAGATTACCTTGTGTCGGGCCCCAACAGACGGGTGCTTACACGAACTCTCGAAGTTC 1120  
QY 1210 ACCAAGTGAATCCAGGAAACATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1269  
Db 1121 ACCAAGTGAATCCAGGAAACATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1180  
QY 1270 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTC 1329  
Db 1181 CCGGCATCCCACTGCTGCAGGAGACAGCCCTGACACTCTCTTTCAGACCCCTCAATCTTC 1240  
QY 1330 CCAGAGATTGAGATGTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGACTCAGG 1389  
Db 1241 CCAGAGATTGAGATGTTCAATCTCTCCAGCCCTGACCCCAATGCTCTCTGACTCAGG 1300  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
Db 1301 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTC 1360  
QY 1450 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCCCTGGGACACTTTTCATCTCAAGC 1509  
Db 1361 CAAAACCTGTCAGGGGGGGTTCGGTCTCAATCTCCCTGGGACACTTTTCATCTCAAGC 1420  
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACT 1565  
Db 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAACT 1476

RESULT 10

US-09-280-116-1  
; Sequence 1, Application US/09280116A  
; Patent No. 6331427  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
; FILE REFERENCE: 5800-24, 035800/176965  
; CURRENT APPLICATION NUMBER: US/09/280,116A  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1504  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: trypsin-like serine proteases  
US-09-280-116-1  
  
Query Match 29.3%; Score 460; DB 3; Length 1504;  
Best Local Similarity 99.6%; Pred. No. 1.2e-113;  
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
  
QY 1090 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCT-G 1148  
Db 1015 CAGGGTGAATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1074  
QY 1149 GGAGATTACCTTGTGTCGGGCCCCAACAGACGGGTGCTTACACGAACTCTCTGCAAGTT 1208



Db	1075	GGGAGATTACCTTGTGCGCGGCCAACAGACGGGTGTCTACACGAACCTCTGCAAGTT	1134
Qy	1209	CACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC	1268
Db	1135	CACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC	1194
Qy	1269	ACGGGCATCCCACTCTCTGCGAGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTT	1328
Db	1195	ACGGGCATCCCACTCTCTGCGAGGACAGCCCTGACACTCTTTTCAGACCCCTCATTCCTT	1254
Qy	1329	CCACAGAGATGTTGAGAATGTTCACTCTCTCAGACCCCTGACCCCAATGTCCTCTGGACTCAG	1388
Db	1255	CCACAGAGATGTTGAGAATGTTCACTCTCTCAGACCCCTGACCCCAATGTCCTCTGGACTCAG	1314
Qy	1389	GGTCTGCTTCCCCACAAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT	1448
Db	1315	GGTCTGCTTCCCCACAAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT	1374
Qy	1449	CCAAAACGTGTCAGGGCGGGGGTTGCCTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAG	1508
Db	1375	CCAAAACGTGTCAGGGCGGGGGTTGCCTCTCAATCTCCCTGGGGCACTTTTCATCCTCAAG	1434
Qy	1509	CTCAGGGCCATCCCTTCTCTCTGACGCTCTGACCCCAAATTTAGT-CCCAGAGAAATAAATCTGA	1567
Db	1435	CTCAGGGCCATCCCTTCTCTCTGACGCTCTGACCCCAAATTTAGTCTCCACGAATTAATCTGA	1494
Qy	1568	GAAG 1571	
Db	1495	GAAG 1498	

RESULT 11  
US-09-949-016-63414/c  
; Sequence 63414, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63414  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63414

Query Match	16.1%; Score 253; DB 4; Length 601;
Best local Similarity	100.0%; Pred. No. 3.5e-58;
Matches 253; Conservative 0;	Mismatches 50; Indels 0; Gaps 0;
QY 1	AGGGAGGTATGGGGAAATTGAAGACAGAGAAACACAAAAATTAGTCCAAAGCGAATGGATTTCTA 60
DB 253	AGGGAGGTATGGGGAAATTGAAGACAGAGAAACACAAAAATTAGTCCAAAGCGAATGGATTTCTA 194
QY 61	TTGGGAGTGATTTTGCCCTTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA 120
DB 193	TTGGGAGTGATTTTGCCCTTAGAAGACACTGGCAATACCAGGAGACATTTTGGTTGTCA 134
QY 121	CAACTATATCGAGGGGCATTACTGGCAACTAAATCGATAGATGCCAAGTGCTGTTTCAAC 180
DB 133	CAACTATATGAGGGGCATTACTGGCAACTAAATCGATAGATGCCAAGTGCTGTTTCAAC 74
QY 181	ATGCTATGATGCAACGGCAGGCCCTCCACAACCAACCAATTATCCAGGCTTCAGATGCCCAAC 240

```

Db      73   ATGCTATGTCACAGCGAGGCCTCCACAACAAACCATTATCCAGTTCAGATGCCCAC    14
|||||
|||

Qy      241  AGTGCCCAGATCG 253
|||||
|||

Db      13  AGTGGCAGATCG 1
|||||
|||

RESULT 12
US-09-949-016-12799
; Sequence 12799, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12799
; LENGTH: 68719
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12799
```

Query Match	14.2%;	Score 223.4;	DB 4;	Length 68719;
Best Local Similarity	86.0%;	Pred. No. 2.8e-49;		
Matches 259;	Conservative 0;	Mismatches 41;	Indels 1;	Gaps 1;
Qy	570	TGAGGCGCAGGCGATGTGGCTCAGCGCTGTAATCCCAACACTTT--GGAGGCGTAGGTCGG	628	
Db	57314	TGAGGCGTGGCCAGCGTGCTTCAAGCTGTAAATCCAGACATTTTGGGAGGCCAAGGCAGG	57373	
Qy	629	CGAATCACTTTGAGGCCAGGAGTTCCAGACACAGCCTGGCCAAACATGGTGAAACCCCGTCTC	688	
Db	57374	TGGATCACCTGAGTTCAGGATTCAGACACACGCTGGCCAAACATGCGGAAACCCCGTCTC	57433	
Qy	689	TACAAAAAAATACAAAAAATTTAGCCGGGTGTGGTGTATGGACACCTGTAGTCAACAGCTAC	748	
Db	57434	TACTAAAAAAATACAAAAAATTTAGCCAGGCATGTGGTGGGCACCTGTAATCCCAGTGAC	57493	
Qy	749	TTGGGAGGCTGAGGCAGGAGAAATTCGTTGAACCCGGGAGATGGAGCTTCAGTGCAGCTGA	808	
Db	57494	TTGGGAGGCTGAGGCAGGAGAAATTCGTTGAACCTGGGAGACAGAGGTTGCGAGTGAGCTGA	57553	
Qy	809	GGTCAGGCCACTTCGCGTCCAACTCTGGGCCAACAGAGCTAAGACTTCCATCTCAAAAAA	868	
Db	57554	GATCGTGCCACTGTACTCCACCTTGGGCGACAGAGAGGACTCCCTCTTCAGAAAAA	57613	
Qy	869	A 869		
Db	57614	A 57614		

RESULT 13  
US-09-949-016-14296  
; Sequence 14296, Application US/09949016  
; Patent No. 681239  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14296
; LENGTH: 68720
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14296

Query Match      14.2%; Score 223.4; DB 4; Length 68720;
Best Local Similarity 86.0%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 570 TTGAGGCCAGGCATGGTGGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 628
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 629 CGAATCACTTGGCCAGGAGTTTCGAGACGAGCTGCGCAACATGGTGAACCCCGTCTC 688
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57374 TGATCACCTGAGTCAAGAGTTCAAGACGAGCTTGCCCAACATGGCGAAACCCCGTCTC 57433
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 689 TACAAAAAATAACAAAAATTAGCCGGTGTGTGATGACACCTGTAGTCACAGCTAC 748
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57434 TACTAAAAAATACAGAAATTAGCCAGGATGGTGGGACCTGTATCCAGTGAC 57493
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 749 TTGGAGGCTGAGCCAGGAGAATTGCTTGAACCCGGGAGATGGAGGCTGCAGTGAGTGA 808
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57494 TTGGAGGCTGAGCCAGGAGAAATTGCTTGAACCTGGGAGACAGAGTTGCGAGTGAGTGA 57553
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 809 GGTGAGGCCACTGGCTCCAACTGGCCACAGAGTAAGACTCCATCTCAAAAAA 868
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57554 GATCGTGCCACTGTACTCCACCTGGGCAACAGAGAGGAGTCCCTCTCAGAAAAA 57613
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 869 A 869
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-949-016-14575
; Sequence 14575, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14575
; LENGTH: 32278
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14575

Query Match      14.2%; Score 222.8; DB 4; Length 32278;
Best Local Similarity 85.8%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 569 TTGAGGCCAGGCATGGTGGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 627
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14296
; LENGTH: 68720
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14296

Query Match      14.2%; Score 223.4; DB 4; Length 68720;
Best Local Similarity 86.0%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 570 TTGAGGCCAGGCATGGTGGCTCAGCGCTGTAATCCCAACACTTTT-GGAGGCTGAGGTGG 628
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 629 CGAATCACTTGGCCAGGAGTTTCGAGACGAGCTGCGCAACATGGTGAACCCCGTCTC 688
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57374 TGATCACCTGAGTCAAGAGTTCAAGACGAGCTTGCCCAACATGGCGAAACCCCGTCTC 57433
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 689 TACAAAAAATAACAAAAATTAGCCGGTGTGTGATGACACCTGTAGTCACAGCTAC 748
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57434 TACTAAAAAATACAGAAATTAGCCAGGATGGTGGGACCTGTATCCAGTGAC 57493
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 749 TTGGAGGCTGAGCCAGGAGAATTGCTTGAACCCGGGAGATGGAGGCTGCAGTGAGTGA 808
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57494 TTGGAGGCTGAGCCAGGAGAAATTGCTTGAACCTGGGAGACAGAGTTGCGAGTGAGTGA 57553
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 809 GGTGAGGCCACTGGCTCCAACTGGCCACAGAGTAAGACTCCATCTCAAAAAA 868
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 57554 GATCGTGCCACTGTACTCCACCTGGGCAACAGAGAGGAGTCCCTCTCAGAAAAA 57613
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 869 A 869
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match      14.2%; Score 222.6; DB 3; Length 176373;
Best Local Similarity 84.0%; Pred. No. 7.2e-49;
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 558 AAAGTTGAACCTTGGAGGCCAGGCATGGTGGCTCAGCGCTGTATATCCCAACACTTTT-GGA 616
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127012 AAAGAAAAAATTATTTGGCCAGGCACGGCGCTCATGCTATATATCCAGCACATTGGGA 127071
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 617 GGCTGAGGTGGGCGAATCACTTGAAGCCAGGAGTTGAGACCCAGGCTGGCCCAACATGGTG 676
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127072 GGCTGAGGAGGTGGATCACTGAGGTTCAGAGTTCGAGACCCAGGCTGGCCCAACATGGTG 127131
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 677 AAACCCCGTCTCTACAAAAAATAACAAAAAATAGCCGGTGTGGTGGTGGACCTGT 736
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127132 AAACCCCATCTCTACTAAAAAATAACAAAAAATAGCCAGGTGTGGTGGGACCTGT 127191
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 737 AGTCACAGCTACTTGGAGGCTCAGGCAGGAGAAATGCTTGAACCCGGGAGATGGAGGCT 796
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127192 AATCCAGCTACACGGGAGGCTCAGGCAGGAGAAATCGCTTGAACCCGGGAGGAGAGTT 127251
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 797 GCAGTGAGGTGAGGTGAGGCCACTGGCGTCAACCTGGGCAACAGAGTAAGACTCCATCT 856
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 127252 GCAGTGAGCCCAAGTTCAAGCCACTGCACCCCGCTGGGCAACAGAGACTTTGTCT 127311

Qy 857 CAAAAAAAAAAAA 869

Db 127312 CAAAAAAAAAAAA 127324

Search completed: July 3, 2005, 18:09:10  
Job time : 181.302 secs

**END PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 628.913 Seconds  
(without alignments)

15667.668 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_10000\_11570

Perfect score: 1571

Sequence: 1 agggaggtatgggaattga.....cccagaataaactgagaag 1571

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq:\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq:\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	541	34.4	567	16	US-10-029-386-7604
2	482	30.7	586	9	US-09-954-531-849
3	482	30.7	586	9	US-09-954-531-1266
4	482	30.7	586	21	US-10-843-641A-1916
5	482	30.7	586	21	US-10-843-641A-2333
6	482	30.7	735	9	US-09-745-288-94
7	482	30.7	735	18	US-10-453-919-94

8	482	30.7	1260	17	US-10-172-118-1411	Sequence 1411, Ap
9	482	30.7	1260	17	US-10-295-027-515	Sequence 515, App
10	482	30.7	1260	17	US-10-173-999-53	Sequence 53, Appl
11	482	30.7	1260	18	US-10-342-887-1411	Sequence 1411, Ap
12	482	30.7	1381	9	US-09-739-907-52	Sequence 52, Appl
13	482	30.7	1381	11	US-09-938-671-52	Sequence 52, Appl
14	482	30.7	1381	11	US-10-935-098-52	Sequence 52, Appl
15	482	30.7	1400	21	US-10-956-157-7041	Sequence 7041, Ap
16	482	30.7	1439	9	US-09-739-907-53	Sequence 53, Appl
17	482	30.7	1439	11	US-09-938-671-53	Sequence 53, Appl
18	482	30.7	1439	11	US-10-935-098-53	Sequence 53, Appl
19	482	30.7	1481	20	US-10-723-860-5032	Sequence 5032, Ap
20	482	30.7	1516	9	US-09-739-907-37	Sequence 37, Appl
21	482	30.7	1516	11	US-09-938-671-37	Sequence 37, Appl
22	482	30.7	1516	21	US-10-935-098-37	Sequence 37, Appl
23	482	30.7	1536	21	US-10-956-157-1806	Sequence 1806, Ap
24	482	30.7	1570	9	US-09-989-722-308	Sequence 308, App
25	482	30.7	1570	9	US-09-989-723-308	Sequence 308, App
26	482	30.7	1570	9	US-09-989-279-308	Sequence 308, App
27	482	30.7	1570	9	US-09-989-727-308	Sequence 308, App
28	482	30.7	1570	9	US-09-989-731-308	Sequence 308, App
29	482	30.7	1570	9	US-09-989-732-308	Sequence 308, App
30	482	30.7	1570	9	US-09-991-073-308	Sequence 308, App
31	482	30.7	1570	9	US-09-990-442-308	Sequence 308, App
32	482	30.7	1570	9	US-09-991-163-308	Sequence 308, App
33	482	30.7	1570	9	US-09-993-604-308	Sequence 308, App
34	482	30.7	1570	9	US-09-990-456-308	Sequence 308, App
35	482	30.7	1570	9	US-09-989-721-308	Sequence 308, App
36	482	30.7	1570	9	US-09-992-598-308	Sequence 308, App
37	482	30.7	1570	9	US-09-989-293A-308	Sequence 308, App
38	482	30.7	1570	9	US-09-989-735-308	Sequence 308, App
39	482	30.7	1570	9	US-09-990-444-308	Sequence 308, App
40	482	30.7	1570	9	US-09-991-181-308	Sequence 308, App
41	482	30.7	1570	9	US-09-989-730-308	Sequence 308, App
42	482	30.7	1570	9	US-09-990-436-308	Sequence 308, App
43	482	30.7	1570	9	US-09-993-687-308	Sequence 308, App
44	482	30.7	1570	10	US-09-989-734-308	Sequence 308, App
45	482	30.7	1570	10	US-09-997-653-308	Sequence 308, App

#### ALIGNMENTS

#### RESULT 1

US-10-029-386-7604

; Sequence 7604, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: ABOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 7604

; LENGTH: 567

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR19.3

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3

; OTHER INFORMATION: EST\_HUMAN HIT: W73140.1, EVALUE 0.00e+00

; OTHER INFORMATION: NT HIT: AF135028.1, EVALUE 0.00e+00

; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUE 6.00e-27

US-10-029-386-7604

```
Query Match      34.4%; Score 541; DB 16; Length 567;
Best Local Similarity 100.0%; Pred. No. 7.6e-150;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1031 GTCTCTTGTGCCACATTGGCCACATCTGCGCTCTCTCATGCCCCCTTTCTCTCTGC 1090
Db 1 GTCTCTTGTGCCACATTGGCCACATCTGCGCTCTCTCATGCCCCCTTTCTCTCTGC 60

Qy 1091 AGGGTGATTTGGGGGCGCTGGTGTGCAATGGCTCCCTCCAGGAGCTGTGCTCTGG 1150
Db 61 AGGGTGATTTGGGGGCGCTGGTGTGCAATGGCTCCCTCCAGGAGCTGTGCTCTGG 120

Qy 1151 GAGATTACCCCTTGTGCCCGGCCCAACAGACGGGTGTCTACAGAACTCTGCAAGTTCA 1210
Db 121 GAGATTACCCCTTGTGCCCGGCCCAACAGACGGGTGTCTACAGAACTCTGCAAGTTCA 180

Qy 1211 CCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 1270
Db 181 CCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 240

Qy 1271 CGGCATCCCACTCTGCGAGGAGACAGCCCTGACACTCTCTTTCAGACCCTCATCTCTCC 1330
Db 241 CGGCATCCCACTCTGCGAGGAGACAGCCCTGACACTCTCTTTCAGACCCTCATCTCTCC 300

Qy 1331 CAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCACTGTCTCTGGACTCAGGG 1390
Db 301 CAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCACTGTCTCTGGACTCAGGG 360

Qy 1391 TCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAACAATTTCC 1450
Db 361 TCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAACAATTTCC 420

Qy 1451 AAAAAGTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGACATTTTCATCTCAAGCT 1510
Db 421 AAAAAGTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGACATTTTCATCTCAAGCT 480

Qy 1511 CAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGAA 1570
Db 481 CAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGAA 540

Qy 1571 G 1571
Db 541 G 541
```

RESULT 2

```
US-09-954-531-849/c
; Sequence 849, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 849
; LENGTH: 586
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-849
```

```
Query Match      30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 CAGGGTGATTTCTGGGGGCGCTGGTGTGCAATGGCTCTGAGGAGCTCGTGTCTCTGG 1149
Db 501 CAGGGTGATTTCTGGGGGCGCTGGTGTGCAATGGCTCTGAGGAGCTCGTGTCTCTGG 442

Qy 1150 GAGATTACCCCTTGTGCCCGGCCCAACAGACGGGTGTCTACAGAACTCTGCAAGTTTC 1209
Db 441 GAGATTACCCCTTGTGCCCGGCCCAACAGACGGGTGTCTACAGAACTCTGCAAGTTTC 382

Qy 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 1269
Db 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCACTCCAGGACTCAGCAC 322

Qy 1270 CCGGCATCCCACTCTGCGAGGAGACAGCCCTGACACTCTCTTTCAGACCCTCATCTCTTC 1329
Db 321 CCGGCATCCCACTCTGCGAGGAGACAGCCCTGACACTCTCTTTCAGACCCTCATCTCTTC 262

Qy 1330 CAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCACTGTCTCTGGACTCAGG 1389
Db 261 CAGAGATGTTGAGATGTTTCACTCTCCAGCCCTGACCCCACTGTCTCTGGACTCAGG 202

Qy 1390 GTCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAACAATTTTC 1449
Db 201 GTCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAACAATTTTC 142

Qy 1450 CAAAAGTCTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGACATTTTCATCTCAAGC 1509
Db 141 CAAAAGTCTCCAGGCGGGGGTGGGTCTCAATCTCCCTGGGGACATTTTCATCTCAAGC 82

Qy 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 1569
Db 81 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 22

Qy 1570 AG 1571
Db 21 AG 20
```

RESULT 3

```
US-09-954-531-1266/c
; Sequence 1266, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1266
; LENGTH: 586
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-1266

Query Match      30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGCTGATTCCTGGGGGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
DB 501 CAGGCTGATTCCTGGGGGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 442
QY 1150 GGAGATTACCTTGTGCTCCCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 441 GGAGATTACCTTGTGCTCCCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
DB 381 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 322
QY 1270 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCCCTTCAGACCCCTCATTCCTTC 1329
DB 321 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCCCTTCAGACCCCTCATTCCTTC 262
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
DB 261 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
DB 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 142
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB 141 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569
DB 81 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 4
US-10-843-641A-1916/c
; Sequence 1916, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/843,641A
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
```

```
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1916
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n=a,t,g or c
US-10-843-641A-1916

Query Match      30.7%; Score 482; DB 21; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGCTGATTCCTGGGGGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 1149
DB 501 CAGGCTGATTCCTGGGGGCTGTGCTCTGCAATGGCTCCCTGCAGGAGCTCGTCTCTGG 442
QY 1150 GGAGATTACCTTGTGCTCCCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 441 GGAGATTACCTTGTGCTCCCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1269
DB 381 ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 322
QY 1270 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCCCTTCAGACCCCTCATTCCTTC 1329
DB 321 CCGGCATCCCCACCTGCTGCAGGACAGCCCTGACACTCCCTTCAGACCCCTCATTCCTTC 262
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
DB 261 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
DB 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC 142
QY 1450 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
DB 141 CAAAACTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 1569
DB 81 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 5
US-10-843-641A-2333/c
; Sequence 2333, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/843,641A
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
```

; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US/09/954,531  
; PRIOR FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/09/954,456  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,436  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/962,832  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/09/964,824  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/09/967,768  
; PRIOR FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/09/968,007  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,347  
; PRIOR FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/09/969,708  
; PRIOR FILING DATE: 2001-10-03  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 8447  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2333  
; LENGTH: 586  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(586)  
; OTHER INFORMATION: n=a,t,g or c  
US-10-843-641A-2333

Query Match 30.7%; Score 482; DB 21; Length 586;  
Best Local Similarity 100.0%; Pred. No. 2.6e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 1149  
DB 501 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 442  
  
QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 1209  
DB 441 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 382  
  
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTGCATCCAGGACTCAGACA 1269  
DB 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTGCATCCAGGACTCAGACA 322  
  
QY 1270 CCGGCATCCCCACCTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCTCTCATTCCTTC 1329  
DB 321 CCGGCATCCCCACCTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCTCTCATTCCTTC 262  
  
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCCTCTGAGCTCAGG 1389  
DB 261 CCAGAGATGTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGTCCTCTGAGCTCAGG 202  
  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
DB 201 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCGGGAACAATTC 142  
  
QY 1450 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509  
DB 141 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82  
  
QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569  
DB 81 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 22  
  
QY 1570 AG 1571  
DB 21 AG 20

RESULT 6  
US-09-745-288-94/c  
; Sequence 94, Application US/09745288  
; Patent No. US20010018058A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.446D1  
; CURRENT APPLICATION NUMBER: US/09/745,288  
; CURRENT FILING DATE: 2000-12-19  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 94  
; LENGTH: 735  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-745-288-94

Query Match 30.7%; Score 482; DB 9; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.8e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 1149  
DB 500 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGGACTCGTGTCTGG 441  
  
QY 1150 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 1209  
DB 440 GGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC 381  
  
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTGCATCCAGGACTCAGACA 1269  
DB 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTGCATCCAGGACTCAGACA 321  
  
QY 1270 CCGGCATCCCCACCTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCTCTCATTCCTTC 1329  
DB 320 CCGGCATCCCCACCTGCTGCGAGGACAGCCCTGACACTCCTTTTCAGACCTCTCATTCCTTC 261  
  
QY 1330 CCAGAGATGTTGAGAAATGTTTCATCTCTCAGGCCCTGACCCCATGTCCTCTGAGCTCAGG 1389  
DB 260 CCAGAGATGTTGAGAAATGTTTCATCTCTCAGGCCCTGACCCCATGTCCTCTGAGCTCAGG 201  
  
QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCGGGAACAATTC 1449  
DB 200 GTCTGCTTCCCCACATTTGGGCTGACCGTCTCTCTAGTTGAACCTCGGGAACAATTC 141  
  
QY 1450 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509  
DB 140 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 81  
  
QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569  
DB 80 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 21  
  
QY 1570 AG 1571  
DB 20 AG 19

RESULT 7  
US-10-453-919-94/c  
; Sequence 94, Application US/10453919  
; Publication No. US2004003230A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY



;; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
;; FILE REFERENCE: 210121.446C7  
;; CURRENT APPLICATION NUMBER: US/10/453,919  
;; CURRENT FILING DATE: 2003-06-03  
;; NUMBER OF SEQ ID NOS: 121  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 94  
;; LENGTH: 735  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
US-10-453-919-94

Query Match 30.7%; Score 482; DB 18; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.8e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 1149  
Db |||||  
QY 500 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 441  
Db |||||  
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209  
Db |||||  
QY 440 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 381  
Db |||||  
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 1269  
Db |||||  
QY 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 321  
Db |||||  
QY 1270 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTTCAGACCTCTATCTCTTC 1329  
Db |||||  
QY 320 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTTCAGACCTCTATCTCTTC 261  
Db |||||  
QY 1330 CCAGAGATGTGAGAAATTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389  
Db |||||  
QY 260 CCAGAGATGTGAGAAATTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 201  
Db |||||  
QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCTGGGAACTTC 1449  
Db |||||  
QY 200 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCTGGGAACTTC 141  
Db |||||  
QY 1450 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCTCCCTGGGACATTTCTATCTCTCAAGC 1509  
Db |||||  
QY 140 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCTCCCTGGGACATTTCTATCTCTCAAGC 81  
Db |||||  
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 1569  
Db |||||  
QY 80 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 21  
Db |||||  
QY 1570 AG 1571  
Db ||  
QY 20 AG 19

RESULT 8  
US-10-172-118-1411  
;; Sequence 1411, Application US/10172118  
;; Publication No. US20030224374A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Dai, Hongyue  
;; APPLICANT: He, Yudong  
;; APPLICANT: Linsley, Peter  
;; APPLICANT: Mao, Mao  
;; APPLICANT: Roberts, Chris  
;; APPLICANT: Van 't Veer, Laura  
;; APPLICANT: Van de Vijver, Marc  
;; APPLICANT: Bernards, Rene  
;; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
;; FILE REFERENCE: 9301-175-999  
;; CURRENT APPLICATION NUMBER: US/10/172,118  
;; CURRENT FILING DATE: 2002-06-14  
;; PRIOR APPLICATION NUMBER: 60/380,770  
;; PRIOR FILING DATE: 2002-05-14  
;; NUMBER OF SEQ ID NOS: 2699

;; SEQ ID NO 1411  
;; LENGTH: 1260  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; PUBLICATION INFORMATION:  
;; DATABASE ACCESSION NUMBER: NM 012427  
;; DATABASE ENTRY DATE: 2001-06-18  
US-10-172-118-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 3.6e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 1149  
Db |||||  
QY 766 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTGG 825  
Db |||||  
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209  
Db |||||  
QY 826 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 885  
Db |||||  
QY 1210 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 1269  
Db |||||  
QY 886 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCTATCCAGGACTCAGACA 945  
Db |||||  
QY 1270 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTTCAGACCTCTATCTCTTC 1329  
Db |||||  
QY 946 CCGGCATCCCACTGCTGCAGGAGACCGCTGACACTCTCTTTTCAGACCTCTATCTCTTC 1005  
Db |||||  
QY 1330 CCAGAGATGTGAGAAATTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389  
Db |||||  
QY 1006 CCAGAGATGTGAGAAATTTCACTCTCCAGCCCTGACCCCATGTCTCTGAGACTCAGG 1065  
Db |||||  
QY 1390 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCTGGGAACTTC 1449  
Db |||||  
QY 1066 GTCTGCTTCCCACTGAGTGGCTGACCGTGTCTCTAGTGAACCTCTGGGAACTTC 1125  
Db |||||  
QY 1450 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCTCCCTGGGACATTTCTATCTCTCAAGC 1509  
Db |||||  
QY 1126 CAAACTGTCCAGGGGGGGTTCGCTCTCAATCTCTCCCTGGGACATTTCTATCTCTCAAGC 1185  
Db |||||  
QY 1510 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 1569  
Db |||||  
QY 1186 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA 1245  
Db |||||  
QY 1570 AG 1571  
Db ||  
QY 1246 AG 1247

RESULT 9  
US-10-295-027-515  
;; Sequence 515, Application US/10295027  
;; Publication No. US20030232350A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Afar, Daniel  
;; APPLICANT: Aziz, Natasha  
;; APPLICANT: Ginsberg, Wendy M.  
;; APPLICANT: Gish, Kurt C.  
;; APPLICANT: Glynn, Richard  
;; APPLICANT: Hevezi, Peter A.  
;; APPLICANT: Mack, David H.  
;; APPLICANT: Murray, Richard  
;; APPLICANT: Watson, Susan R.  
;; APPLICANT: Eos Biotechnology, Inc.  
;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
;; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
;; FILE REFERENCE: 018501-012500US  
;; CURRENT APPLICATION NUMBER: US/10/295,027  
;; CURRENT FILING DATE: 2002-11-13  
;; PRIOR APPLICATION NUMBER: US 09/663,733  
;; PRIOR FILING DATE: 2000-09-15  
;; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 515  
; LENGTH: 1260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-295-027-515

Query Match 30.7%; Score 482; DB 17; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 3.6e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1090	CAGGGTGATTCGCGGGGCTGTGGTCTGCAATGGCTCCCTGAGGACTCGTGTCTGG	1149
DB	766	CAGGGTGATTCGCGGGGCTGTGGTCTGCAATGGCTCCCTGAGGACTCGTGTCTGG	825
QY	1150	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC	1209
DB	826	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC	885
QY	1210	ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA	1269
DB	886	ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA	945
QY	1270	CGGCAATCCCACTGCTGAGGAGCAGCGCTGACACTCTTTCAGACCTCATCTCTTC	1329
DB	946	CGGCAATCCCACTGCTGAGGAGCAGCGCTGACACTCTTTCAGACCTCATCTCTTC	1005
QY	1330	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGCACTCAGG	1389
DB	1006	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGCACTCAGG	1065
QY	1390	GTCTGCTTCCCACTGAGCGTGTCTCTAGTTGAACCCCTGGGAACAATTC	1449
DB	1066	GTCTGCTTCCCACTGAGCGTGTCTCTAGTTGAACCCCTGGGAACAATTC	1125
QY	1450	CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC	1509
DB	1126	CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC	1185
QY	1510	TCAGGCGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAAATAAATACTGAGA	1569
DB	1186	TCAGGCGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAAATAAATACTGAGA	1245
QY	1570	AG 1571	
DB	1246	AG 1247	

RESULT 10  
US-10-173-999-53  
; Sequence 53, Application US/10173999  
; Publication No. US20040005563A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002420US  
; CURRENT APPLICATION NUMBER: US/10/173,999  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/299,234  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: US 60/315,287  
; PRIOR FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2001-04-12  
; NUMBER OF SEQ ID NOS: 163  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 1260  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-173-999-53

Query Match 30.7%; Score 482; DB 17; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 3.6e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1090	CAGGGTGATTCGCGGGGCTGTGGTCTGCAATGGCTCCCTGAGGACTCGTGTCTGG	1149
DB	766	CAGGGTGATTCGCGGGGCTGTGGTCTGCAATGGCTCCCTGAGGACTCGTGTCTGG	825
QY	1150	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC	1209
DB	826	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTGCAAGTTC	885
QY	1210	ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA	1269
DB	886	ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCAATCCAGGACTCAGCACA	945
QY	1270	CGGCAATCCCACTGCTGAGGAGCAGCGCTGACACTCTTTCAGACCTCATCTCTTC	1329
DB	946	CGGCAATCCCACTGCTGAGGAGCAGCGCTGACACTCTTTCAGACCTCATCTCTTC	1005
QY	1330	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGCACTCAGG	1389
DB	1006	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGTCTCTGCACTCAGG	1065
QY	1390	GTCTGCTTCCCACTGAGCGTGTCTCTAGTTGAACCCCTGGGAACAATTC	1449
DB	1066	GTCTGCTTCCCACTGAGCGTGTCTCTAGTTGAACCCCTGGGAACAATTC	1125
QY	1450	CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC	1509
DB	1126	CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC	1185
QY	1510	TCAGGCGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAAATAAATACTGAGA	1569
DB	1186	TCAGGCGCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCCAAAATAAATACTGAGA	1245
QY	1570	AG 1571	
DB	1246	AG 1247	

RESULT 11  
US-10-342-887-1411  
; Sequence 1411, Application US/10342887  
; Publication No. US20040058340A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.

APPLICANT: Mao, Mao  
APPLICANT: Roberts, Christopher J.  
APPLICANT: Van 't Veer, Laura Johanna  
APPLICANT: Van de Vijver, Marc J.  
APPLICANT: Bernards, Rene  
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
FILE REFERENCE: 9101-188-999  
CURRENT APPLICATION NUMBER: US/10/342,887  
CURRENT FILING DATE: 2003-01-15  
PRIOR APPLICATION NUMBER: 60/298,918  
PRIOR FILING DATE: 2001-06-18  
PRIOR APPLICATION NUMBER: 60/380,710  
PRIOR FILING DATE: 2002-05-14  
PRIOR APPLICATION NUMBER: 10/172,118  
PRIOR FILING DATE: 2002-06-14  
NUMBER OF SEQ ID NOS: 2699  
SEQ ID NO 1411  
LENGTH: 1260  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-342-887-1411

Query Match 30.7%; Score 482; DB 18; Length 1260;  
Best Local Similarity 100.0%; Pred. No. 3.6e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG	1149
Db	766	CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG	825
Qy	1150	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTTCGCAAGTTC	1209
Db	826	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTTCGCAAGTTC	885
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTAGTTCATCCAGGACTCAGCACA	1269
Db	886	ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTAGTTCATCCAGGACTCAGCACA	945
Qy	1270	CCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC	1329
Db	946	CCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC	1005
Qy	1330	CCAGAGATGTTGAGAAATGTTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1389
Db	1006	CCAGAGATGTTGAGAAATGTTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1065
Qy	1390	GTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1449
Db	1066	GTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1125
Qy	1450	CAAACTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1509
Db	1126	CAAACTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1185
Qy	1510	TCAGGGCCATCCCTTCTCTGACGCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569
Db	1186	TCAGGGCCATCCCTTCTCTGACGCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1245
Qy	1570	AG 1571	
Db	1246	AG 1247	

RESULT 12  
US-09-739-907-52  
Sequence 52, Application US/09739907  
Patent No. US20010012889A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 36 Human Secreted Proteins  
FILE REFERENCE: P2022P1  
CURRENT APPLICATION NUMBER: US/09/739,907  
CURRENT FILING DATE: 2000-12-20

PRIOR APPLICATION NUMBER: 09/348,457  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: 60/070,567  
PRIOR FILING DATE: 1998-01-07  
PRIOR APPLICATION NUMBER: 60/070,692  
PRIOR FILING DATE: 1998-01-07  
PRIOR APPLICATION NUMBER: 60/070,704  
PRIOR FILING DATE: 1998-01-07  
PRIOR APPLICATION NUMBER: 60/070,658  
PRIOR FILING DATE: 1998-01-07  
NUMBER OF SEQ ID NOS: 196  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 52  
LENGTH: 1381  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-739-907-52

Query Match 30.7%; Score 482; DB 9; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 3.7e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG	1149
Db	852	CAGGGTGAATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG	911
Qy	1150	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTTCGCAAGTTC	1209
Db	912	GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTTCGCAAGTTC	971
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTAGTTCATCCAGGACTCAGCACA	1269
Db	972	ACCAAGTGGATCCAGGAAACCATCCAGGCCAATCTCTAGTTCATCCAGGACTCAGCACA	1031
Qy	1270	CCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC	1329
Db	1032	CCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCTCTTTCAGACCCCTCATTCCTTC	1091
Qy	1330	CCAGAGATGTTGAGAAATGTTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1389
Db	1092	CCAGAGATGTTGAGAAATGTTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1151
Qy	1390	GTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1449
Db	1152	GTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAAACAATTC	1211
Qy	1450	CAAACTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1509
Db	1212	CAAACTGTCCAGGGCGGGGTTCGGTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC	1271
Qy	1510	TCAGGGCCATCCCTTCTCTGACGCTTGCAGCTTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569
Db	1272	TCAGGGCCATCCCTTCTCTGACGCTTGCAGCTTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1331
Qy	1570	AG 1571	
Db	1332	AG 1333	

RESULT 13  
US-09-938-671-52  
Sequence 52, Application US/09938671  
Publication No. US2004002066A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 36 Human Secreted Proteins  
FILE REFERENCE: P2022P1  
CURRENT APPLICATION NUMBER: US/09/938,671  
CURRENT FILING DATE: 2001-08-27  
PRIOR APPLICATION NUMBER: 09/348,457  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: 60/070,567  
PRIOR FILING DATE: 1998-01-07

```
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-52

Query Match      30.7%; Score 482; DB 11; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 1149
Db      |||
852 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 911
QY 1150 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db      |||
912 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 971
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTGCATCCAGGACTCAGCACA 1269
Db      |||
972 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTGCATCCAGGACTCAGCACA 1031
QY 1270 CCGGCATCCCACTCTGTCAGGGAGACGCCCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1329
Db      |||
1032 CCGGCATCCCACTCTGTCAGGGAGACGCCCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1091
QY 1330 CCAGAGATGTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1389
Db      |||
1092 CCAGAGATGTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1151
QY 1390 GTCTGCTTCCCACTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1449
Db      |||
1152 GTCTGCTTCCCACTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1211
QY 1450 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 1509
Db      |||
1212 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 1271
QY 1510 TCAGGGCCCACTCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1569
Db      |||
1272 TCAGGGCCCACTCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1331
QY 1570 AG 1571
Db      ||
1332 AG 1333
```

```
RESULT 14
US-10-935-098-52
; Sequence 52, Application US/10935098
; Publication No. US20050042667A1
; GENERAL INFORMATION:
; APPLICANT: Lafleur et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022p1c3
; CURRENT APPLICATION NUMBER: US/10/935,098
; CURRENT FILING DATE: 2004-09-08
; PRIOR APPLICATION NUMBER: 09/938,671
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/739,907
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: PCT/US99/00108
; PRIOR FILING DATE: 1999-01-06
```

```
; PRIOR APPLICATION NUMBER: 60/070,657
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-935-098-52

Query Match      30.7%; Score 482; DB 21; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 1149
Db      |||
852 CAGGGTGATTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 911
QY 1150 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db      |||
912 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 971
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTGCATCCAGGACTCAGCACA 1269
Db      |||
972 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTGCATCCAGGACTCAGCACA 1031
QY 1270 CCGGCATCCCACTCTGTCAGGGAGACGCCCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1329
Db      |||
1032 CCGGCATCCCACTCTGTCAGGGAGACGCCCTGACACTCTCTTTTCAGACCCCTCATTCCTTC 1091
QY 1330 CCAGAGATGTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1389
Db      |||
1092 CCAGAGATGTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1151
QY 1390 GTCTGCTTCCCACTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1449
Db      |||
1152 GTCTGCTTCCCACTGAGATGTTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGAGTGCATCCAGG 1211
QY 1450 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 1509
Db      |||
1212 CAAAACCTGTCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCCTCAAGC 1271
QY 1510 TCAGGGCCCACTCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1569
Db      |||
1272 TCAGGGCCCACTCCCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCCAAGAAATAAACTGAGA 1331
QY 1570 AG 1571
Db      ||
1332 AG 1333
```

```
RESULT 15
US-10-956-157-7041
; Sequence 7041, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7041
; LENGTH: 1400
```

TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-956-157-7041

Query Match 30.7%; Score 482; DB 21; Length 1400;  
Best Local Similarity 100.0%; Pred. No. 3.7e-132;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1090	CAGGGTGATTTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG	1149
Db	881	CAGGGTGATTTCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG	940
Qy	1150	GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC	1209
Db	941	GGAGATTACCTTGTGCCCCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAGTTC	1000
Qy	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA	1269
Db	1001	ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAATCCAGGACTCAGCACA	1060
Qy	1270	CCGGCATCCCACTGTCTGAGGACAGCCCTGACACTCCTTTCAGACCCCTCATTCCTTC	1329
Db	1061	CCGGCATCCCACTGTCTGAGGACAGCCCTGACACTCCTTTCAGACCCCTCATTCCTTC	1120
Qy	1330	CCAGAGATGTTGAGAAATGTTCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1389
Db	1121	CCAGAGATGTTGAGAAATGTTCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG	1180
Qy	1390	GTCTGCTTCCCCCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC	1449
Db	1181	GTCTGCTTCCCCCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC	1240
Qy	1450	CAAAACTGTCCAGGGCGGGGTTCCGTCCTCAATCTCCCTGGGGCACTTTCATCTCAAGC	1509
Db	1241	CAAAACTGTCCAGGGCGGGGTTCCGTCCTCAATCTCCCTGGGGCACTTTCATCTCAAGC	1300
Qy	1510	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1569
Db	1301	TCAGGGCCCATCCCTTCTCTGCACTCTGACCCAAATTTAGTCCAGAAATAAATCTGAGA	1360
Qy	1570	AG 1571	
Db	1361	AG 1362	

Search completed: July 4, 2005, 02:13:09  
Job time : 630.913 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 3415.96 Seconds  
(without alignments)  
17505.766 Million cell updates/sec

Title: US-09-936-271c-13\_COPY\_10000\_11570

Perfect score: 1571

Sequence: 1 agggagtgatgggaattga.....cccagaataaactgagaag 1571

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	482	30.7	539	2	BF513278 UI-H-BW1-
C 2	482	30.7	586	7	W73140 zd55e11.s1
C 3	475.8	30.3	809	4	BG680075 602628224
C 4	455	29.0	478	1	AA862032 0146e09.s
C 5	443.8	28.2	802	4	BG682309 602629626
C 6	440	28.0	953	2	BE745465 601579834
C 7	431.8	27.5	666	4	BM842155 K-EST0119
C 8	417	26.5	645	1	AI002163 0t42h05.s
C 9	383.6	24.4	677	2	BE898385 601681219
C 10	381.8	24.3	453	7	N80762 za98f06.s1
C 11	355.8	22.6	363	2	AW105502 xd53g06.x
C 12	348	22.2	590	4	BM840511 K-EST0117
C 13	333.4	21.2	577	4	BM838406 K-EST0114
C 14	322.8	20.5	430	7	W68361 zd34h08.s1
C 15	317	20.2	388	7	W68496 zd34h08.r1
C 16	314	20.0	334	2	BF514439 UI-H-BW1-
C 17	291	18.5	533	4	BM840509 K-EST0117
C 18	278.8	17.7	523	4	BM841293 K-EST0118
C 19	277.2	17.6	320	2	AW380655 RC4-HT027
C 20	276.6	17.6	411	2	AW801647 ILS-UM006
C 21	273	17.4	512	4	BM830263 K-EST0103
C 22	270.4	17.2	526	4	BM841697 K-EST0118
C 23	245	15.6	498	7	W73168 zd55e11.r1
C 24	229.8	14.6	420	6	CB298043 220013_re

## ALIGNMENTS

RESULT 1  
BF513278/c

LOCUS

DEFINITION UI-H-BW1-amo-e-03-0-UI.s1 NCI CGAP\_Sub7 Homo sapiens cdna clone  
IMAGE:3070564 3', mRNA sequence.

ACCESSION BF513278

VERSION BF513278.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 539)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-x@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:

NCI-CGAP clone distribution information can be found through the

I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. 539

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3070564"

/lab\_host="NCI CGAP Sub7"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7

is a subtracted library derived from NCI CGAP Sub6. The

single-stranded DNA preparation of NCI CGAP Sub6 was used

as a tracer in a subtractive hybridization with a driver

comprising the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM

3334-3337, 3682-3683, 3798-3803 IMAGE Clones

1322376-1323911, 1456008-1456775, 1500552-1502855);

NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778

(IMAGE Clones 1323912-1325831, 1471368-1472903,

25	229.2	14.6	422	4	BG398353
26	227.8	14.5	947	4	BG680848
27	227.6	14.5	479	2	AW576251
28	227.6	14.5	486	1	AI755214
29	227.6	14.5	724	5	BQ005946
30	227.6	14.5	726	7	BN395303
31	227.6	14.5	961	5	BU197099
32	227.6	14.5	4544	3	BSM803689
33	226	14.4	359	2	AW500684
34	226	14.4	476	1	AI754567
35	226	14.4	481	1	AI754105
36	226	14.4	740	7	BN395308
37	226	14.4	977	6	CD516130
38	226	14.4	1092	4	BM480074
39	226	14.4	1547	3	CR621085
40	226	14.4	1672	3	CR615950
41	223.4	14.2	3341	3	CR749847
42	223	14.2	744	4	BG743962
43	222.4	14.2	715	2	BF574982
44	222	14.1	352	1	AI249688
45	221.2	14.1	649	4	BG939464

BF513278 539 bp mRNA linear EST 07-DEC-2000

UI-H-BW1-amo-e-03-0-UI.s1 NCI CGAP\_Sub7 Homo sapiens cdna clone

IMAGE:3070564 3', mRNA sequence.

BF513278

BF513278.1 GI:11598457

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:

NCI-CGAP clone distribution information can be found through the

I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. 539

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:3070564"

/lab\_host="NCI CGAP Sub7"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; NCI CGAP Sub7

is a subtracted library derived from NCI CGAP Sub6. The

single-stranded DNA preparation of NCI CGAP Sub6 was used

as a tracer in a subtractive hybridization with a driver

comprising the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM

3334-3337, 3682-3683, 3798-3803 IMAGE Clones

1322376-1323911, 1456008-1456775, 1500552-1502855);

NCI CGAP Kids pool 1 LLAM 3338-3342, 3722-3725, 3776-3778

(IMAGE Clones 1323912-1325831, 1471368-1472903,

1492104-1493255); NCI\_CGAP\_Lu5\_pool 1 LLAM 3575-3582, 3851-3854 (IMAGE ClonEIDs 1414920-1417991, 1520904-1522439); NCI\_CGAP\_GC4\_pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE ClonEIDs 1257096-1258631, 1459064-1470983, 1475592-1476743); NCI\_CGAP\_Pr22\_pool 1 LLAM 2457-2459, 2759-2759, 3062-3068 (IMAGE ClonEIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI\_CGAP\_Co10\_pool 1 LLAM 2644-2653, 2871-2872 (IMAGE ClonEIDs 1057416-1061255, 1144584-1145351). (6% of the driver population), plus a pool of 3,840 arrayed clones from NCI\_CGAP\_Sub1 (IMAGE ClonEIDs 2708616-2710535) and NCI\_CGAP\_Sub2 (IMAGE ClonEIDs 2710536-27121455) (4% of the driver population), plus a pool of 11,136 clones from NCI\_CGAP\_Sub3 (IMAGE ClonEIDs 2712456-2723391) (10% of the driver population), plus a pool of 5,472 clones from NCI\_CGAP\_Sub4 (IMAGE ClonEIDs 2723592-2729326) (40% of the driver population), plus a pool of 4032 clones from NCI\_CGAP\_Sub5 (IMAGE ClonEIDs 2728969-2733190) (40% of the driver population). Subtraction was performed as previously described (Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806. TAG\_TISSUE=lung TAG\_LIB=NCI\_CGAP-Lu13 TAG\_SEQ=GCCGG"

## ORIGIN

Query Match	30.7%;	Score 482;	DB 2;	Length 539;
Best Local Similarity	100.0%;	Pred. No. 4.8e-65;		
Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1090	CAGGGTCATTCTGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	1149	
Db	508	CAGGGTCATTCTGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	449	
QY	1150	GGAGATTACCTTTGTGCGCGGCCCAACAGACCCGGGGTGTCTACACGAACCTCTGCAAGTTC	1209	
Db	448	GGAGATTACCTTTGTGCGCGGCCCAACAGACCCGGGGTGTCTACACGAACCTCTGCAAGTTC	389	
QY	1210	ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTCTAGTTCATCCAGGACTCAGGACA	1269	
Db	388	ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTCTAGTTCATCCAGGACTCAGGACA	329	
QY	1270	CCGSCATCCCCACCTGCTGCAGGAGACAGCCCTGCACACTCTCTTTTCAGACCCCTCAATTCCTTC	1329	
Db	328	CCGSCATCCCCACCTGCTGCAGGAGACAGCCCTGCACACTCTCTTTTCAGACCCCTCAATTCCTTC	269	
QY	1330	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGACACTCAGG	1389	
Db	268	CCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCAATGTCCTCTGACACTCAGG	209	
QY	1390	GTCTGCTTCCCCCACAATTTGGGCTGACCGTGTCTCTTAGTTGAACCTCTGGGAACAATTTTC	1449	
Db	208	GTCTGCTTCCCCCACAATTTGGGCTGACCGTGTCTCTTAGTTGAACCTCTGGGAACAATTTTC	149	
QY	1450	CAAAACTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC	1509	
Db	148	CAAAACTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC	89	
QY	1510	TCAGGGGCCATCCCTTCTCTGCACTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA	1569	
Db	88	TCAGGGGCCATCCCTTCTCTGCACTCTGACCCCAAAATTTAGTCCAGAAATAAAGTGA	29	
QY	1570	AG 1571		
Db	28	AG 27		
RESULT 2				
LOCUS	W73140/c			
DEFINITION	zds5ell.s1 Soares fetal heart N8H1.9W Homo sapiens cDNA clone			
LOCUS	W73140	586 bp	mRNA	linear
EST	16-OCT-1996			

RESULT 2  
W73140/C

LOCUS	W73140	586 bp	mRNA	linear	EST 16-OCT-1996
DEFINITION	zds5ell.s1	Soares fetal heart	NbHH19W	Homo sapiens	cDNA clone

IMAGE:344588 3' similar to PIR:A53968 A53968 serine proteinase SCCE  
precursor - human ;, mRNA sequence.

ACCESSION W73140  
VERSION W73140.1 GI:1383275  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 586)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,  
Parsons,J., Rifkin,L., Rohlffing,T., Soares,M., Tan,P.,  
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and  
Wilson,R.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert length: 761 Std Error: 0.00  
Seq primer: ETPrimer  
High quality sequence stop: 428.

FEATURES  
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1269963"
/db_xref="taxon:9606"
/clone="IMAGE:344588"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_heart_NDHL19W"
/notes="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCACTCTCAAGTCGAGCGCGGCATCTTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Patima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NDHL19W."

```

## ORIGIN

	Query Match	30.7%;	Score 482;	DB 7;	Length 586;	
	Best Local Similarity	100.0%;	Pred. No. 4.6e-65;			
	Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1090	CAGGGTGATTCTCGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG	1149			
Db	501	CAGGGTGATTCTCGGGGGCCTGTGGTCTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG	442			
QY	1150	GGAGATTACCTTTGTGCCCGGCCCAAACAGACCGGGTGTCTACAGAACTCTTGCAGATTTC	1209			
Db	441	GGAGATTACCTTTGTGCCCGGCCCAAACAGACCGGGTGTCTACAGAACTCTTGCAGATTTC	382			
QY	1210	ACCAATGGATTCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGACATCAGACA	1269			
Db	381	ACCAATGGATTCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGACATCAGACA	322			
QY	1270	CGGGCATCCCCACTCTGCGAGGACAGCCCTGACACTCTTTTCAGACCTCATTCCTTC	1329			
Db	321	CGGGCATCCCCACTCTGCGAGGACAGCCCTGACACTCTTTTCAGACCTCATTCCTTC	262			



```

Qy 1330 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1389
Db 261 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 202
Qy 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 1449
Db 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAACAATTC 142
Qy 1450 CAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 1509
Db 141 CAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 82
Qy 1510 TCAGGGCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACTGAGA 1569
Db 81 TCAGGGCCATCCCTTCTCTCAGCTCTGACCCCAAAATTTAGTCCAGAAATAAACTGAGA 22
Qy 1570 AG 1571
Db 21 AG 20

RESULT 3
LOCUS BG680075 809 bp mRNA linear EST 01-MAY-2001
DEFINITION 602628224F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753372 5',
mRNA sequence.
ACCESSION BG680075
VERSION BG680075.1 GI:13911472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10612 row: o column: 05
High quality sequence stop: 707.
FEATURES
source
1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753372"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCWV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5Kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 30.3%; Score 475.8; DB 4; Length 809;
Best Local Similarity 99.6%; Pred. No. 3.8e-64;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1084 CTCCTGAGGTGATCTGGGGGCGCTGTGTCTGCAATGCTCCCTGACGGGACTCGTG 1143
Db 219 CTCCTGAGGTGATCTGGGGGCGCTGTGTCTGCAATGCTCCCTGACGGGACTCGTG 278
Qy 1144 TCCTGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACAGAACTCTGC 1203

```

```

Db 279 TCCTGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACAGAACTCTGC 338
Qy 1204 AAGTTCCCAAGTGATGCCAGGAAACCATCCAGGCCAACTCTCTGAGTCATCCCAAGGACTC 1263
Db 339 AAGTTCCCAAGTGATGCCAGGAAACCATCCAGGCCAACTCTCTGAGTCATCCCAAGGACTC 398
Qy 1264 AGCACACGGCATCCCACTGCTGACGGGACAGCCCTGACACTCCCTTCAGACCCCTCAT 1323
Db 399 AGCACACGGCATCCCACTGCTGACGGGACAGCCCTGACACTCCCTTCAGACCCCTCAT 458
Qy 1324 TCCTTCCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGA 1383
Db 459 TCCTTCCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGA 518
Qy 1384 CTGAGGTCGTCTTCCCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAC 1443
Db 519 CTCAGGGTCGTCTTCCCCCAACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAC 578
Qy 1444 AATTTCCAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCC 1503
Db 579 AATTTCCAAACTGTCCAGGGCGGGGTTGGGTCTCAATCTCCCTGGGGCACTTTCATCC 638
Qy 1504 TCAAGCTCAGGGCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAA 1562
Db 639 TCAAGCTCAGGGCCATCCCTTCTCTCAGCTCTGACCCAAATTTAGTCCAGAAATAA 697

RESULT 4
LOCUS AA862032 478 bp mRNA linear EST 13-APR-1999
DEFINITION O146809 s1 NCI_CGAP_HN3 Homo sapiens cDNA clone IMAGE:1485736 3',
similar to SW:SCCE_HUMAN P49862 STRATUM CORNEUM CHYMOTRIPTIC ENZYME
PRECUSOR ; mRNA sequence.
ACCESSION AA862032
VERSION AA862032.1 GI:2954511
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 478)
AUTHORS NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: John Ensley, M.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1647 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 233.
FEATURES
source
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1485736"
/tissue_type="squamous cell carcinoma from base of tongue"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP_HN3"
/notes="Organ: tongue; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.0 kb. 5' adaptor sequence:
5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5'
(GA)10ACTAGTCTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN

```

Query Match 29.0%; Score 455; DB 1; Length 478;  
Best Local Similarity 99.8%; Pred. No. 7.5e-61;  
Matches 466; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1105 GGGCTGTGTCTGCAATGCTCTCCAGGGACTGTGTCCTGGGGAGATTACCCCTTGT 1164  
Db 478 GGGCTGTGTCTGCAATGCTCTCCAGGGACTGTGTCCTGGGGAGATTACCCCTTGT 420  
QY 1165 GCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCCACCAAGTGGATCCAG 1224  
Db 419 GCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTCCACCAAGTGGATCCAG 360  
QY 1225 GAAACCATCCAGGCCAACTCTCAGTCAATCCAGGACTCAGACACCGGATCCCCACCT 1284  
Db 359 GAAACCATCCAGGCCAACTCTCAGTCAATCCAGGACTCAGACACCGGATCCCCACCT 300  
QY 1285 GCTGAGGGACAGCCCTGACATCTCTTTTCAGACCTCATTTCTCCAGAGATGTTGAGA 1344  
Db 299 GCTGAGGGACAGCCCTGACATCTCTTTTCAGACCTCATTTCTCCAGAGATGTTGAGA 240  
QY 1345 ATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGGGTCTGTTCCCCAC 1404  
Db 239 ATGTTATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGGGTCTGTTCCCCAC 180  
QY 1405 ATGCGGCTGACCGTGTCTCTAGTTGAACCTCGGAAACAAATTTCCAAAACCTGTCCAGGG 1464  
Db 179 ATGCGGCTGACCGTGTCTCTAGTTGAACCTCGGAAACAAATTTCCAAAACCTGTCCAGGG 120  
QY 1465 CGGGGGTGGCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGCCATCCCT 1524  
Db 119 CGGGGGTGGCTCAATCTCCCTGGGGCACTTTTCATCTCAAGCTCAGGGCCATCCCT 60  
QY 1525 TCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGAG 1571  
Db 59 TCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATCTGAGAG 13

RESULT 5  
BG682309 802 bp mRNA linear EST 01-MAY-2001  
LOCUS 602629626F1 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',  
DEFINITION mRNA sequence.  
ACCESSION BG682309.1 GI:13913706  
VERSION BG682309.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 802)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10615 row: d column: 08  
High quality sequence stop: 802.  
Location/Qualifiers  
1. .802  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:4754263"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Skn4"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
Query Match 28.2%; Score 443.8; DB 4; Length 802;  
Best Local Similarity 99.3%; Pred. No. 3.5e-59;  
Matches 456; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 1090 CAGGGTGATTCGGGGGGCTGTGTGTCGCAATGGCTCCCTCGAGGACTCGTGTCTCTGG 1149  
Db 333 CAGGGTGATTCGGGGGGCTGTGTGTCGCAATGGCTCCCTCGAGGACTCGTGTCTCTGG 392  
QY 1150 GGAGATTACCTTGTCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209  
Db 393 GGAGATTACCTTGTCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 452  
QY 1210 ACCAAGTGGATCCAGGAAACCAATCCAGGCAACCTCTCTGAGTCAATCCAGGACTCAGCACA 1269  
Db 453 ACCAAGTGGATCCAGGAAACCAATCCAGGCAACCTCTCTGAGTCAATCCAGGACTCAGCACA 512  
QY 1270 CGGGCATCCCACTCTGTCAGGAGCAGCCCTGACACTCTTTTCAGACCCCTCAITCTCTTC 1329  
Db 513 CGGGCATCCCACTCTGTCAGGAGCAGCCCTGACACTCTCTTTTCAGACCCCTCAITCTCTTC 572  
QY 1330 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCATGTCCTCTGAGTCAATCCAGG 1389  
Db 573 CCAGAGATGTTGAGAAATGTTTCACTCTCCAGCCCTGACCCCATGTCCTCTGAGTCAATCCAGG 632  
QY 1390 GTCTGTTCTCCCCACACA-TTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT 1448  
Db 633 GTCTGTTCTCCCCACACA-TTGGGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAACAATTT 692  
QY 1449 CCAAAACTGTCCAGGGGGGGTGGCTCTCAATCTCCCTGGGGGACATTTTCATCTCTCAAG 1508  
Db 693 CCAAAACTGTCCAGGGGGGGTGGCTCTCAATCTCCCTGGGGGACATTTTCATCTCTCAAG 752  
QY 1509 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTT 1547  
Db 753 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTT 791

RESULT 6  
BE745465 953 bp mRNA linear EST 15-SEP-2000  
LOCUS 601579834F1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3928640 5',  
DEFINITION mRNA sequence.  
ACCESSION BE745465  
VERSION BE745465.1 GI:10159457  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 953)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCW760 row: c column: 09  
High quality sequence stop: 685.  
Location/Qualifiers  
1. .953  
/organism="Homo sapiens"  
/mol\_type="mRNA"

```

/db_xref="taxon:9606"
/clone="IMAGE:3928640"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/notes="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

## ORIGIN

```

Query Match      28.0%; Score 440; DB 2; Length 953;
Best Local Similarity 99.6%; Pred. No. 1.3e-58;
Matches 462; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1090 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCGCAGGGACTCGTCTCTGG 1149
Db 229 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCGCAGGGACTCGTCTCTGG 288
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1209
Db 289 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 348
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCAC 1269
Db 349 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCAC 408
QY 1270 CCGGCATCCCATCTGTCGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329
Db 409 CCGGCATCCCATCTGTCGAGGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 468
QY 1330 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1389
Db 469 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 528
QY 1390 GTCTGCTTCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 1449
Db 529 GTCTGCTTCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 587
QY 1450 CAAAAGTGTCCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 1509
Db 588 CAAAAGTGTCCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 647
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCC 1553
Db 648 TCAGGGCCCATCCCTTCTCTGACG-TCTGACCCCAATTTAGTCC 690
```

```

RESULT 7
BM842155 666 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0119498 S12SNU216 Homo sapiens cDNA clone S12SNU216-56-H11 5',
DEFINITION mRNA sequence.
ACCESSION BM842155
VERSION BM842155.1 GI:19198564
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 666)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
CONTACT Contact: Kim YS
COMMENT Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
```

```

52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 56 row: H column: 11
High quality sequence stop: 666.
Location/Qualifiers
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-56-H11"
/sex="np"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S12SNU216"
```

## FEATURES

```

Source
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-56-H11"
/sex="np"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S12SNU216"
/notes="Organ: Stomach; Vector: PCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then dephosphorylated
with tobacco acid pyrophosphatase (TAP). The dephosphorylated
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

## ORIGIN

```

Query Match      27.5%; Score 431.8; DB 4; Length 666;
Best Local Similarity 99.5%; Pred. No. 2.7e-57;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGGACTCGTCTCTGG 1149
Db 232 CAGGGTGATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGGACTCGTCTCTGG 291
QY 1150 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209
Db 292 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 351
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCAC 1269
Db 352 ACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCTCTGAGTTCATCCAGGACTCAGCAC 411
QY 1270 CCGGCATCCCATCTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329
Db 412 CCGGCATCCCATCTGCTGAGGAGACAGCCCTGACACTCTCTTTCAGACCTTCATCTCTTC 471
QY 1330 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1389
Db 472 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 531
QY 1390 GTCTGCTTCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 1449
Db 532 GTCTGCTTCCGCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAAACCTCGGAAACAATTC 591
QY 1450 CAAAAGTGTCCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 1509
Db 592 CAAAAGTGTCCAGGGCGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC 651
QY 1510 TCAGGGCCCATCCCT 1524
Db 652 TCAGGGCCCATCCCT 666
```

```

RESULT 8
AI002163/c
LOCUS
DEFINITION
  Ot-42H05.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1619481
  3', similar to TR:Q61955 Q61955 SERINE PROTEASE INHIBITOR 5 ; , mRNA
  sequence.
ACCESSION
  AI002163
VERSION
  1
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 645)
  NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-r@mail.nih.gov
  CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
  Bonaldo, Ph.D.
  CDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CCGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 1042 Std Error: 0.00
  Seq primer: -40ml3 fwd. ET from Amersham
  High quality sequence stop: 303.
FEATURES
  source
    1..645
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:1619481"
    /sex="male"
    /lab_host="DH10B"
    /clone_lib="Soares testis NHT"
    /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
    was prepared from mRNA obtained from Clontech
    Laboratories, Inc., and primed with a Not I - oligo(dT)
    primer [5',
    TGTACCACTCTGAAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
    Double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I
    and Eco RI sites of the modified pT7T3 vector. Library
    went through one round of normalization to Cot5, and was
    constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
  Query Match 26.5%; Score 417; DB 1; Length 645;
  Best Local Similarity 94.9%; Pred. No. 5.3e-55;
  Matches 464; Conservative 0; Mismatches 20; Indels 5; Gaps 3;
  QY 1084 CTCTGCAGGCTGATTCTGGGGGCTGTGTCTGCAATGCTGCTCCGAGGACTCGTG 1143
  Db 499 CTCTGCAGGCTGATTCTGGGGGCTGTGTCTGCAATGCTGCTCCGAGGCTCGTGT- 441
  QY 1144 TCCTGGGGAGATTACCTCTGTGCCCGCCCAACAGACCGGGGTGTCTACAGAACTCTGC 1203
  Db 440 --CCTGGGAGATTACCTTGT-CCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGC 384
  QY 1204 AGTTTCAACAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCATCCAGGACTC 1263
  Db 383 AAGTTTCAACAGTGGATCCAGGAACCAATCCAGGCCAATCTCTGAGTCATCCAGGACTC 324
  QY 1264 AGCACACCGGCATCCCACTGCTGTCAGGAGCAGCCCTGACACTCCTTTCCAGACCCCTCAT 1323
  Db 323 AGCACACCGGCATCCCACTGCTGTCAGGAGCAGCCCTGACACTCCTTTCCAGACCCCTCAT 264
  QY 1324 TCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCCTGGA 1383

```

```

Db 263 TCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCCTGGA 204
QY 1384 CTCAGGCTCTGCTTCCCTCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCTTGGGAAC 1443
Db 203 CTCAGGCTCTGCTTCCCTCCACATTTGGGTGACCGTGTCTCTCTAGTTGAACCTTGGGAAC 144
QY 1444 AATTTCCAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCC 1503
Db 143 AATTTCCAAAACCTGTCCAGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCC 84
QY 1504 TCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCGACCCCAAAATTTAGT-CCCAGAAATAA 1562
Db 83 TCAAGCTCAGGGCCCATCCCTTCTCTGACGCTCGACCCCAAAATTTAGTCCCAGAAATAA 24
QY 1563 ACTGAGAAG 1571
Db 23 ACTGAGAAG 15
RESULT 9
BE989385
LOCUS
DEFINITION
  601691219F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951365 S',
  mRNA sequence.
ACCESSION
  BE989385
VERSION
  BE989385.1 GI:10364809
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 677)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-r@mail.nih.gov
  Tissue Procurement: DCTD/DTp
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
  Plate: LLCW819 row: f column: 06
  High quality sequence stop: 677.
FEATURES
  source
    1..677
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:3951365"
    /tissue_type="adenocarcinoma cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH MGC 9"
    /notes="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
    EcoRI; cDNA made by oligo-dT priming. Directionally
    cloned into EcoRI/XhoI sites using the following 5'
    adaptor: GGCACGAG(G). Size-selected >500bp for average
    insert size 1.8kb. Library constructed by Ling Hong in
    the laboratory of Gerald M. Rubin (University of
    California, Berkeley) using ZAP-cDNA synthesis kit
    (Stratagene) and Superscript II RT (Life Technologies). "
ORIGIN
  Query Match 24.4%; Score 383.6; DB 2; Length 677;
  Best Local Similarity 98.5%; Pred. No. 7.8e-50;
  Matches 397; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
  QY 1090 CAGGGTGATTCTGGGGGGCTGTGCTGCAATGGCTCCCTGAGGACTCGTGTCTGG 1149
  Db 276 CAGGGTGATTCTGGGGGGCTGTGCTGCAATGGCTCCCTGAGGACTCGTGTCTGG 335

```

QY 1150 GGAGATTACCTTGTGCGCCGCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 1209  
|||||  
Db 336 GGAGATTACCTTGTGCGCCGCGCCCAACAGACCGGGTGTCTACAGAACCTCTGCAAGTTC 395  
|||||  
QY 1210 ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 1269  
|||||  
Db 396 ACCAAGTGGATCCAGGAAACCATCCAGGCCCAACTCTCTGAGTTCATCCAGGACTCAGCACA 455  
|||||  
QY 1270 CCGGCATCCCACTGTCTGAGGAGACCGCTGACACTCTCTTTCAGACCTTCATCTCTTC 1329  
|||||  
Db 456 CCGGCATCCCACTGTCTGAGGAGACCGCTGACACTCTCTTTCAGACCTTCATCTCTTC 515  
|||||  
QY 1330 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTTCAGG 1389  
|||||  
Db 516 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGAGTTCAGG 575  
|||||  
QY 1390 GTCTGTTCCCCCACTGAGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAATTC 1449  
|||||  
Db 576 GNTCTGTTCCCCCACTGAGGCTGACCGTGTCTCTCTAGTTGAACCCCTGGGAAATTC 635  
|||||  
QY 1450 CAAACTGTCCAGGCGGGGGTTCGGTCTCAATCTCCCTGGGG 1492  
|||||  
Db 636 CAAACTGTCCA-GCGCGGGGGTTCGGTCTCAATCTCCCTGGGG 677  
|||||

## RESULT 10

N80762/c  
LOCUS  
DEFINITION N80762 453 bp mRNA linear EST 29-MAR-1996  
za98f06.s1 Soares fetal lung NbHL19W Homo sapiens cDNA clone  
IMAGE:300611 3', mRNA sequence.

ACCESSION N80762.1 GI:1243463

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 453)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,

Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,

Parsons, J., Rifkin, L., Rohlfs, T., Soares, M., Tan, F.,

Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and

Wilson, R.

The WashU-Merck EST Project

Unpublished (1995)

Contact: Willson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: eat@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: ml3 -40 forward

High quality sequence stop: 319.

## FEATURES

Location/Qualifiers

1..453

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:124535"

/db\_xref="taxon:9606"

/clone="IMAGE:300611"

/dev\_stage="19 weeks"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares fetal lung NbHL19W"

/note="Organ: lung; Vector: pF73D (Pharmacia) with a

modified polylinker; Site: 1: Not 1; Site 2: Eco RI; 1st

strand cDNA was primed with a Not 1 - oligo(dT) primer

[5'-TGTAACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'],

double-stranded cDNA was size selected, ligated to Eco RI

adapters (Pharmacia), digested with Not 1 and cloned into

the Not 1 and Eco RI sites of a modified pF73 vector

(Pharmacia). Library went through one round of

normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. This library was constructed  
from the same fetus as the fetal heart library. Soares  
fetal heart NbHL19W."

## ORIGIN

Query Match 24.3%; Score 381.8; DB 7; Length 453;  
Best Local Similarity 96.6%; Pred. No. 1.7e-49;  
Matches 422; Conservative 0; Mismatches 9; Indels 6; Gaps 3;

QY 1141 GTGTCTCTGGGAGATTACCTTGTGCGCGCCCAACAGAGAA---CCGGGTGTCTACAGAA 1196  
|||||

Db 441 GTGTCTTGGNAGATTACCTTGTGCGCGCCCAACAGAGCGGGGTGTCTACAGAA 382  
|||||

QY 1197 CTTCTGCAAGTTCACC-AAGTGGATCCAGGAAACCATCCAGGCGCAACTCTCTGAGTTC 1255  
|||||

Db 381 CTTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCGCAACTCTCTGAGTTC 322  
|||||

QY 1256 CAGGACTCAGCACACCGGCATCCCACTGC-TGCAGGGACAGCCCTGACACTCTTTCA 1314  
|||||

Db 321 CAGGACTCAGCACACCGGCATCCCACTGC-TGCAGGGACAGCCCTGACACTCTTTCA 262  
|||||

QY 1315 GACCTTCATTCTCTCCAGAGATGTTGAGAAATGTTTCATCTCTCAGCCCTGACCCCATG 1374  
|||||

Db 261 GACCTTCATTCTCTCCAGAGATGTTGAGAAATGTTTCATCTCTCAGCCCTGACCCCATG 202  
|||||

QY 1375 TCTCTGAGTTCAGGGTCTCTCTCCCAACATTCAGGCGTCTCTCTAGTGAAC 1434  
|||||

Db 201 TCTCTGAGTTCAGGGTCTCTCTCCCAACATTCAGGCGTCTCTCTAGTGAAC 142  
|||||

QY 1435 CTTGGGAAACAATTTCCAAAACCTGTCAGGGGGGGTTCGCTCTCAATCTCCCTGGGGCA 1494  
|||||

Db 141 CTTGGGAAACAATTTCCAAAACCTGTCAGGGGGGGTTCGCTCTCAATCTCCCTGGGGCA 82  
|||||

QY 1495 CTTTCATCTCTCAAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCC 1554  
|||||

Db 81 CTTTCATCTCTCAAGCTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCC 22  
|||||

QY 1555 AGAAATAAATGAGAG 1571  
|||||

Db 21 AGAAATAAATGAGAG 5  
|||||

## RESULT 11

AW105502/c

LOCUS

DEFINITION

AW105502

363 bp mRNA linear EST 20-OCT-1999

xd53g06.x1 NCI\_CGAP\_Ov23 Homo sapiens cDNA clone IMAGE:2597530 3',

mRNA sequence.

ACCESSION

AW105502.1

GI:6076237

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 363)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Possible reversed clone: polyT not found

Seq primer: -40UP from Gibco

```
FEATURES
  source
    High quality sequence stop: 362.
    Location/Qualifiers
      1..363
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2597530"
        /tissue_type="tumor, 5 pooled (see description)"
        /lab_host="DH10B"
        /clone_lib="NCI CGAP Ov23"
        /notes="Organ: Ovary; Vector: pCMV-SPORT6; Site 1: SalI;
        Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 1.35 kb. Tumor types include: mixed
        Mullerian tumor, papillary serous, clear cell, spindle
        cell. All are primary tumors, metastasis positive. Life
        Technologies catalog #: 11534-013"

ORIGIN
Query Match      22.6%; Score 355.8; DB 2; Length 363;
Best Local Similarity 99.4%; Pred. No. 1.9e-45;
Matches 357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1213 AAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACACCG 1272
Db      |||||
QY 363 AAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACACCG 304
Db      |||||
QY 1273 GCATCCCACTGCTGCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTCCCA 1332
Db      |||||
QY 303 GCATCCCACTGCTGCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTCCCA 244
Db      |||||
QY 1333 GAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGGGTC 1392
Db      |||||
QY 243 GAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCCTCTGGACTCAGGGTC 184
Db      |||||
QY 1393 TGCTTCCCACTAGGGCTGACGGTGTCTCTAGTTGAACCTCGGGAACAATTTCCAA 1452
Db      |||||
QY 183 TGCTTCCCACTAGGGCTGACGGTGTCTCTAGTTGAACCTCGGGAACAATTTCCAA 124
Db      |||||
QY 1453 AACTGTCCAGGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTTCATCTCAAGTCA 1512
Db      |||||
QY 123 AACTGTCCAGGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTTCATCTCAAGTCA 64
Db      |||||
QY 1513 GGCCCACTCCTCTCTGCGAGCTGTGACCCAAATTTAGTCCCAAGAAATAAAGTGAAG 1571
Db      |||||
QY 63 GGCCCACTCCTCTCTGCGAGCTGTGACCCAAATTTAGTCCCAAGAAATAAAGTGAAG 5

RESULT 12
BM840511 BM840511 590 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0117576 S12SNU216 Homo sapiens cDNA clone S12SNU216-39-G08 5',
DEFINITION mRNA sequence.
ACCESSION BM840511.1 GI:19196920
VERSION BM840511.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 590)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsun@mail.kribb.re.kr
Plate: 39 row: G column: 08
```

```
FEATURES
  source
    High quality sequence stop: 590.
    Location/Qualifiers
      1..590
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="S12SNU216-39-G08"
        /sex="F"
        /tissue_type="Lymph node"
        /cell_type="Epithelial"
        /cell_line="SNU-216"
        /lab_host="Top10F"
        /clone_lib="S12SNU216"
        /notes="Organ: Stomach; Vector: pcNS; Site 1: EcoRI;
        Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
        bacterial alkaline phosphatase (BAP) and then decapped
        with tobacco acid pyrophosphatase (TAP). The decapped
        intact mRNA was ligated with DNA-RNA linker including EcoR
        I site by treatment of T4 RNA ligase and the first strand
        cDNA was synthesized from oligo dt-selected mRNA by
        priming with dt-tailed vector. The dt-tailed vector was
        adjusted to have about 60nt. The cDNA vector was
        circularized with E. coli DNA ligase after digestion of
        EcoRI which site is also included in vector. An RNA strand
        converted to a DNA strand by Okayama-Berg method. The
        obtained cDNA vectors were used for transfection of
        competent cells E. coli Top10F' by electroporation method.
        The cDNA libraries constructed by this method are
        full-length enriched cDNA library."
```

ORIGIN

Query Match 22.2%; Score 348; DB 4; Length 590;  
Best Local Similarity 100.0%; Pred. No. 2.7e-44;  
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGAGACTCGTGTCTCTGG 1149  
Db |||||

QY 243 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGAGACTCGTGTCTCTGG 302  
Db |||||

QY 1150 GGNAGATTACCTTGTCGCCGGCCCAACAGACGGGGTCTACAGAACCTTCGAAGTTC 1209  
Db |||||

QY 303 GGNAGATTACCTTGTCGCCGGCCCAACAGACGGGGTCTACAGAACCTTCGAAGTTC 362  
Db |||||

QY 1210 ACCAAGTGATCCAGGAACCATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACA 1269  
Db |||||

QY 363 ACCAAGTGATCCAGGAACCATCCAGGCCAACTCCTGAGTCATCCAGGACTCAGCACA 422  
Db |||||

QY 1270 CCGGCATCCCACTCTCTCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCAATCTTC 1329  
Db |||||

QY 423 CCGGCATCCCACTCTCTCAGGAGCAGCCCTGACACTCCTTTTCAGACCCCTCAATCTTC 482  
Db |||||

QY 1330 CCAGAGATCTTGAGATGTTTCATCTCTCAGCCCTGAGCCCATGCTCTCTGGACTCAGG 1389  
Db |||||

QY 483 CCAGAGATCTTGAGATGTTTCATCTCTCAGCCCTGAGCCCATGCTCTCTGGACTCAGG 542  
Db |||||

QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCT 1437  
Db |||||

QY 543 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCT 590  
Db |||||

RESULT 13  
BM838406 BM838406 577 bp mRNA linear EST 06-MAR-2002  
LOCUS K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5',  
DEFINITION mRNA sequence.  
ACCESSION BM838406  
VERSION BM838406.1 GI:19194815  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 577)











QY	241	ACACAAATGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCAGATACACGC	300
Db	241	ACACAAATGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCAGATACACGC	300
QY	301	AAAGGGCAGAGACACAGTTTTTCAGGGTGTGTCTATGATCATCTCTTTTTTTTTTTTT	360
Db	301	AAAGGGCAGAGACACAGTTTTTCAGGGTGTGTCTATGATCATCTCTTTTTTTTTTTTT	360
QY	361	TTTTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCGG	420
Db	361	TTTTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCGG	420
QY	421	GATCTCGGCTCACTGCAAGCTCCGGCTCCGGGTTCACGCCATTCTCTCCCTCAGCCTC	480
Db	421	GATCTCGGCTCACTGCAAGCTCCGGCTCCGGGTTCACGCCATTCTCTCCCTCAGCCTC	480
QY	481	CCAAGTAGCTGGAGTACAGGGCGCCGCCACTACGCCCGGCTAAATTTTTTGTATTTTA	540
Db	481	CCAAGTAGCTGGAGTACAGGGCGCCGCCACTACGCCCGGCTAAATTTTTTGTATTTTA	540
QY	541	GTAGAGACGGGGTTTCAACCGTTTTAGCCGGGATGGCTCGATCTCTGACCTCGTGATCC	600
Db	541	GTAGAGACGGGGTTTCAACCGTTTTAGCCGGGATGGCTCGATCTCTGACCTCGTGATCC	600
QY	601	GCCCGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGAAGCCACCGCGCCCGGCCATG	660
Db	601	GCCCGCTCGGCTCCCAAGTGTGGGATTTACAGGCGTGAAGCCACCGCGCCCGGCCATG	660
QY	661	ATCATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
Db	661	ATCATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	720
QY	721	AATATGAGTTTGGGCGAGCCCGTGGCTCATGCTGATGATGATGATGATGATGATGATG	780
Db	721	AATATGAGTTTGGGCGAGCCCGTGGCTCATGCTGATGATGATGATGATGATGATGATG	780
QY	781	AGAGTGGGTGAATCACTTGAGGCGCAGGAGTTTGAGACACAGCTGGCCAAATGGTGAA	840
Db	781	AGAGTGGGTGAATCACTTGAGGCGCAGGAGTTTGAGACACAGCTGGCCAAATGGTGAA	840
QY	841	CTCTGTCTTTTACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	900
Db	841	CTCTGTCTTTTACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	900
QY	901	ACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCTCGAGGCG	960
Db	901	ACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCTCGAGGCG	960
QY	961	GAGGTTGACGTGGCGCGAGATCAATCACCGCCCTCCAGGCTGGGCGACAGAGCAAGACT	1020
Db	961	GAGGTTGACGTGGCGCGAGATCAATCACCGCCCTCCAGGCTGGGCGACAGAGCAAGACT	1020
QY	1021	CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1080
Db	1021	CTGTCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1080
QY	1081	AAAAAATAATGCTCAACAAATAGACAGAGTGAATTAAGGAAATAAATAAATAAATAA	1140
Db	1081	AAAAAATAATGCTCAACAAATAGACAGAGTGAATTAAGGAAATAAATAAATAAATAA	1140
QY	1141	AGAACTTAAGGTATATTTGACAAATCATTTAGAACTTTTAAATAAATAAATAAATAA	1200
Db	1141	AGAACTTAAGGTATATTTGACAAATCATTTAGAACTTTTAAATAAATAAATAAATAA	1200
QY	1201	GGCATGAAAGACAGGGAGGAAACAGGGAGACAGAAACACTGTGGGCCCAAGAGACAA	1260
Db	1201	GGCATGAAAGACAGGGAGGAAACAGGGAGACAGAAACACTGTGGGCCCAAGAGACAA	1260
QY	1261	ACAAGGCTCTTACAGACAGAGGAGGAGAGAGAGTGTGAGTGTGAGAGACAGAGAG	1320
Db	1261	ACAAGGCTCTTACAGACAGAGGAGGAGAGAGAGTGTGAGTGTGAGAGACAGAGAG	1320

QY	1321	AAAAACACAGAGAGAGAGACACAGACACAGAGAGACAGAGAGCGGAGAGGATAGAAAGA	1380
Db	1321	AAAAACACAGAGAGAGAGACACAGACACAGAGAGACAGAGAGCGGAGAGGATAGAAAGA	1380
QY	1381	GAGAGAGGGGTGGAGAGACACACAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
Db	1381	GAGAGAGGGGTGGAGAGACACACAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
QY	1441	GAACACACAGAGAGATGGAAGAGACTCTCAGAAAAAACACAGACAAAAAGATGAAAGAGG	1500
Db	1441	GAACACACAGAGAGATGGAAGAGACTCTCAGAAAAAACACAGACAAAAAGATGAAAGAGG	1500
QY	1501	AGTATCGAGGGTGAACACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
Db	1501	AGTATCGAGGGTGAACACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
QY	1561	TCCAGGCGGCAAGATAGTAGACCCAGAGTTGGTGAAGCCAGATCCTTAAAGCTGGGGG	1620
Db	1561	TCCAGGCGGCAAGATAGTAGACCCAGAGTTGGTGAAGCCAGATCCTTAAAGCTGGGGG	1620
QY	1621	AGSCAGGGAAGGGCTGGCTTCCGAGACCCCTCCCAATTTCTCCGGGCCAGGGAG	1680
Db	1621	AGSCAGGGAAGGGCTGGCTTCCGAGACCCCTCCCAATTTCTCCGGGCCAGGGAG	1680
QY	1681	GTAGGAGTGAATTTCCGACCTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGGAGCA	1740
Db	1681	GTAGGAGTGAATTTCCGACCTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGGAGCA	1740
QY	1741	GGAGGAGCTATTGCTAAGGCCGATAGGACACTCATTTGCCGGGAATGTGCCCAGGGAG	1800
Db	1741	GGAGGAGCTATTGCTAAGGCCGATAGGACACTCATTTGCCGGGAATGTGCCCAGGGAG	1800
QY	1801	CAGTGGTGTATTAATCTCAGGCCCGTCCAGAGCCAGAGAGGAGGAGTGCACAGGA	1860
Db	1801	CAGTGGTGTATTAATCTCAGGCCCGTCCAGAGCCAGAGAGGAGGAGTGCACAGGA	1860
QY	1861	AGSCACAGGCTTGAGAGTCTCGGCTGAGCTGGGAGCAAAATCCCAACCCCTACCTGG	1920
Db	1861	AGSCACAGGCTTGAGAGTCTCGGCTGAGCTGGGAGCAAAATCCCAACCCCTACCTGG	1920
QY	1921	GGGACAGGCAAGTGAAGCTGGTGAAGGTGGCTCAGAGGCAAGGAGAGAGTGTCT	1980
Db	1921	GGGACAGGCAAGTGAAGCTGGTGAAGGTGGCTCAGAGGCAAGGAGAGAGTGTCT	1980
QY	1981	GTGCGTCTGCAACCATCTTCTCTGCTCCCTCTCTGCTCCCTCTGAGAGGCTGTAGA	2040
Db	1981	GTGCGTCTGCAACCATCTTCTCTGCTCCCTCTCTGCTCCCTCTGAGAGGCTGTAGA	2040
QY	2041	CTCCTATCTTGAATTTCTATAGTCTGGGTCTCAGCGCAGTGGCGCCGTC	2100
Db	2041	CTCCTATCTTGAATTTCTATAGTCTGGGTCTCAGCGCAGTGGCGCCGTC	2100
QY	2101	CTTGTGGTCTCTCTACCTGGGAAATAAGGTAGGGAGGAGGGGAAGTGGGTTAAGG	2160
Db	2101	CTTGTGGTCTCTCTACCTGGGAAATAAGGTAGGGAGGAGGGGAAGTGGGTTAAGG	2160
QY	2161	GCTCCCGGATGCTGGGCTCCCAACCTCTGACATTTCCCAATCCAGGTGACGCGGCC	2220
Db	2161	GCTCCCGGATGCTGGGCTCCCAACCTCTGACATTTCCCAATCCAGGTGACGCGGCC	2220
QY	2221	ATGGCTACAGCAAGACCCCTCTGGATGTGGGTCTCTGTGCTCTGATCAAGCTTGTCTT	2280
Db	2221	ATGGCTACAGCAAGACCCCTCTGGATGTGGGTCTCTGTGCTCTGATCAAGCTTGTCTT	2280
QY	2281	CTGGGGTCAAGGTAAACAGAACTCTGGGTGGGAGGGTGTGGGATTTGGGAGGACTGT	2340
Db	2281	CTGGGGTCAAGGTAAACAGAACTCTGGGTGGGAGGGTGTGGGATTTGGGAGGACTGT	2340
QY	2341	CTCTGGGCACTAGAGGCTGTCCCTTGGGAACTGTGTGAGCTGGGCACTACTCCGG	2400
Db	2341	CTCTGGGCACTAGAGGCTGTCCCTTGGGAACTGTGTGAGCTGGGCACTACTCCGG	2400
QY	2401	GACCGGGTGAATGTGAGTCTCTGTGTACTTGTGGTGTGCGATCGTATGTGCGCCCTGT	2460



```
QY 4621 CTCTTATTTCTCAGGCCCTGCCCTGCTCCCTCAGCATGTCAGACACCCACCTCTAGCT 4680
Db 4621 CTCTTATTTCTCAGGCCCTGCCCTGCTCCCTCAGCATGTCAGACACCCACCTCTAGCT 4680
QY 4681 GGTCTGGCTCTTGTAGTCTGAACACCCACCCAGCCAGCCGCTCTGAGCCCGCC 4740
Db 4681 GGTCTGGCTCTTGTAGTCTGAACACCCACCCAGCCAGCCGCTCTGAGCCCGCC 4740
QY 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCAAACATGATGTTTCTGTGACACC 4800
Db 4741 CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCAAACATGATGTTTCTGTGACACC 4800
QY 4801 CCTCTAACACCTGCTCTGGAGCAACACGAGACCTGGAGCTGGGGCCGGGGAAGACG 4860
Db 4801 CCTCTAACACCTGCTCTGGAGCAACACGAGACCTGGAGCTGGGGCCGGGGAAGACG 4860
QY 4861 CCGGTGCGATGACAGCAGAGCCGCGATCAATGATCCGATCGGATATGACACCC 4920
Db 4861 CCGGTGCGATGACAGCAGAGCCGCGATCAATGATCCGATCGGATATGACACCC 4920
QY 4921 AGCCGTGGAGCCGCGCTGCTTAAGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
Db 4921 AGCCGTGGAGCCGCGCTGCTTAAGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980
QY 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 2
CQ874885
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent WO2004077060.
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE 1
AUTHORS Diamandis, E.P. and Petraki, C.
TITLE Assay for detection of renal cell carcinoma
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;
Mount Sinai Hospital (CA)
FEATURES
source
1. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCAGAGTGAAGCAAGAGAGGAGTTGAGAGCTCCCTCTGCAAGTGGCTTGATC 60
Db 1 GGGCCAGAGTGAAGCAAGAGAGGAGTTGAGAGCTCCCTCTGCAAGTGGCTTGATC 60
QY 61 TCCCTTGCCTTAAATGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTTGCCTTAAATGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 181 ACAGAGAGCCTGGGACACACAGGACACACAGAGTCAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCCTGGGACACACAGGACACACAGAGTCAGAGAGAGAGAGAGAGAGAG 240
```

```
QY 241 ACACAAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACG 300
Db 241 ACACAAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCAGATACACG 300
QY 301 AAAGGGGAGAGACACAGTTTTTCAGGGTGGTGTCTATGATCATCTCTTTTTTTTTTTT 360
Db 301 AAAGGGGAGAGACACAGTTTTTCAGGGTGGTGTCTATGATCATCTCTTTTTTTTTTTT 360
QY 361 TTTTTTTTTTTTTTTTCAGAGCGAGTCTCGCTCTGTGCTCCAGGCTGGAGTGCAGTGGCG 420
Db 361 TTTTTTTTTTTTTTTTCAGAGCGAGTCTCGCTCTGTGCTCCAGGCTGGAGTGCAGTGGCG 420
QY 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCCATTCTCTCTCTCAGCCTC 480
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCCATTCTCTCTCTCAGCCTC 480
QY 481 CCAAGTAGCTGGGACTACAGGGCCCGCCACTACAGCCCGGCTAAATTTTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGGACTACAGGGCCCGCCACTACAGCCCGGCTAAATTTTTTGTATTTTA 540
QY 541 GTAGAGACGGGTTTCAACCGTTTTAGCCGGGATGGCTCGATCTCCTGACCTCGTGATCC 600
Db 541 GTAGAGACGGGTTTCAACCGTTTTAGCCGGGATGGCTCGATCTCCTGACCTCGTGATCC 600
QY 601 GCCCGCTCGGCTCCCAAAGTCTCGGATTAACAGCGTGAGCCACCGCGCCCGGCCATG 660
Db 601 GCCCGCTCGGCTCCCAAAGTCTCGGATTAACAGCGTGAGCCACCGCGCCCGGCCATG 660
QY 661 ATCATCTTCTGACTATGCTGATGTCACAGTACCTTAAGCCATCAGACTCTACCTTTTA 720
Db 661 ATCATCTTCTGACTATGCTGATGTCACAGTACCTTAAGCCATCAGACTCTACCTTTTA 720
QY 721 AATATCAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAATCCAGACCTTTGGGAGGC 780
Db 721 AATATCAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAATCCAGACCTTTGGGAGGC 780
QY 781 AGAGTGGGTGAATCACTTTGAGGCCAGGAGTTTGAGACCAAGCTCGGCCAAACATGGTAAA 840
Db 781 AGAGTGGGTGAATCACTTTGAGGCCAGGAGTTTGAGACCAAGCTCGGCCAAACATGGTAAA 840
QY 841 CTCGTCTTTTACTAAAAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCGTGGGCGAC 900
Db 841 CTCGTCTTTTACTAAAAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCGTGGGCGAC 900
QY 901 ACCTGTAATCCAGCTATGCTGAGGCTGAGGACAGAGTCACTTGAACCTCGAGGCG 960
Db 901 ACCTGTAATCCAGCTATGCTGAGGCTGAGGACAGAGTCACTTGAACCTCGAGGCG 960
QY 961 GAGTTGCAAGTGGGCGGAGATCATACCGCCCTCCAGCTGGGCGACAGAGCAAGACT 1020
Db 961 GAGTTGCAAGTGGGCGGAGATCATACCGCCCTCCAGCTGGGCGACAGAGCAAGACT 1020
QY 1021 CTGTCTCAATTAATAATAACAAACGACAGCAGTTTGTGTACCTTAGTTATATCT 1080
Db 1021 CTGTCTCAATTAATAATAACAAACGACAGCAGTTTGTGTACCTTAGTTATATCT 1080
QY 1081 AAAAAAATAATGCTGTCAACAAATAGACAGAGTGAATAAAGGAAAAATAATGGGCCA 1140
Db 1081 AAAAAAATAATGCTGTCAACAAATAGACAGAGTGAATAAAGGAAAAATAATGGGCCA 1140
QY 1141 AGAATCTTAAGGTATATTTGACAAATCATTTAGAACCTTTAAAAAAGAAAGATACAGA 1200
Db 1141 AGAATCTTAAGGTATATTTGACAAATCATTTAGAACCTTTAAAAAAGAAAGATACAGA 1200
QY 1201 GGCATAGAAAGACAGGGAGGAAACAGGGAGACAGAAACACCTGTGGCCCAAGGAAACAA 1260
Db 1201 GGCATAGAAAGACAGGGAGGAAACAGGGAGACAGAAACACCTGTGGCCCAAGGAAACAA 1260
QY 1261 ACAAGGCTCTTAAGACACAGCAGGAGGAGAGAGAGAGTGTAGTGTAGAGACAGACAG 1320
Db 1261 ACAAGGCTCTTAAGACACAGCAGGAGGAGAGAGAGAGTGTAGTGTAGAGACAGACAG 1320
QY 1321 AAAAAACAGAGAGAGAGAGACAGACAGAGAGACAGAGAGCGGAGAGGATAGAAAGA 1380
```



```
Qy 3541 TCGGCGCATTTGCATCTCAGACCTGGGCAACAGAGCGGAAACTCCGCTCTCGAAGAAAAAAA 3600
Db 3541 TCGGCGCATTTGCATCTCAGACCTGGGCAACAGAGCGGAAACTCCGCTCTCGAAGAAAAAAA 3600
Qy 3601 GAAAAAAGGCTAAGAACAGTGAATGGGACGGGAGGACTGATGATGGAGTGGGC 3660
Db 3601 GAAAAAAGGCTAAGAACAGTGAATGGGACGGGAGGACTGATGATGGAGTGGGC 3660
Qy 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGCAT 3720
Db 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGCAT 3720
Qy 3721 GTTTTCACTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGCCCCCTTAAG 3780
Db 3721 GTTTTCACTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGCCCCCTTAAG 3780
Qy 3781 CTGAGCCCTTCTTCTGCTGCTGCTTTCGAAACCTTAGCTCCGCCCATGAGTCTGACC 3840
Db 3781 CTGAGCCCTTCTTCTGCTGCTGCTTTCGAAACCTTAGCTCCGCCCATGAGTCTGACC 3840
Qy 3841 CCACCTCTTTCCTCAACAGCGCCCTAGGCGCAGACTCTAGTGAGACCCCGCCTAAGGCCA 3900
Db 3841 CCACCTCTTTCCTCAACAGCGCCCTAGGCGCAGACTCTAGTGAGACCCCGCCTAAGGCCA 3900
Qy 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGGTACTTCTTAGAACCCCTTCAA 3960
Db 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGGTACTTCTTAGAACCCCTTCAA 3960
Qy 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTCCAGGCTGGAG 4020
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTCCAGGCTGGAG 4020
Qy 4021 TGCAGTGGCGTGAATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
Db 4021 TGCAGTGGCGTGAATCTGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080
Qy 4081 GCCTCCACCTCCTGAGTAGCTGGGATTACAGGTGCGGCGCACCAACCGCTGGCTAAATTTT 4140
Db 4081 GCCTCCACCTCCTGAGTAGCTGGGATTACAGGTGCGGCGCACCAACCGCTGGCTAAATTTT 4140
Qy 4141 GTGTCTTTAGTAGAGACAGGTTTTCACCTTTGGGCGAGGCTGTCTCAAACTCCCAACC 4200
Db 4141 GTGTCTTTAGTAGAGACAGGTTTTCACCTTTGGGCGAGGCTGTCTCAAACTCCCAACC 4200
Qy 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTCTGGGTTACAGGCGTGAGCCACGC 4260
Db 4201 TCAGGTGATCGGCCACTCGGCTCCAGAGTCTGGGTTACAGGCGTGAGCCACGC 4260
Qy 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGAGACTCTAAACATGTAACCTGACCTGGCC 4320
Db 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGAGACTCTAAACATGTAACCTGACCTGGCC 4320
Qy 4321 CTAACTAAGTCAATTCAAACCCCTCTCGCTCCAGCTGACCCCTGACCTGAGGCC 4380
Db 4321 CTAACTAAGTCAATTCAAACCCCTCTCGCTCCAGCTGACCCCTGACCTGAGGCC 4380
Qy 4381 TGACCCCACTCTTGAGACAGTTCATCCCTAAGCCCTGGTCTCCCTCCATCCCCAG 4440
Db 4381 TGACCCCACTCTTTGAGACAGTTCATCCCTAAGCCCTGGTCTCCCTCCATCCCCAG 4440
Qy 4441 GCTCCAGCCCAAGTGGGACTACCCCTGAGCTTGTCAGGAATCTGTACCCCAAT 4500
Db 4441 GCTCCAGCCCAAGTGGGACTACCCCTGAGCTTGTCAGGAATCTGTACCCCAAT 4500
Qy 4501 TTTACCTCAGATGATTTAGCCAAATTCAGGAATCTGTAGGTCAGTTAGAGTCCA 4560
Db 4501 TTTACCTCAGATGATTTAGCCAAATTCAGGAATCTGTAGGTCAGTTAGAGTCCA 4560
Qy 4561 GTAAACCTTACCTGAGCTGGCTGTCTGCTTGAAGCTTGAAGCTGGCTTGAAGTGGCA 4620
Db 4561 GTAAACCTTACCTGAGCTGGCTGTCTGCTTGAAGCTTGAAGCTGGCTTGAAGTGGCA 4620
```

```
Qy 4621 CTCTTATTTCTCAGGCGCTGCCCCCTGCCCCCTCAGCATGTCAGACACCCACCCCTCTAGCT 4680
Db 4621 CTCTTATTTCTCAGGCGCTGCCCCCTGCCCCCTCAGCATGTCAGACACCCACCCCTCTAGCT 4680
Qy 4681 GGTCTGGCCTTTGAGTCTGAAACCCACCCCGAGCCCAAGCCCGGCTCTGAGCCCCGCC 4740
Db 4681 GGTCTGGCCTTTGAGTCTGAAACCCACCCCGAGCCCAAGCCCGGCTCTGAGCCCCGCC 4740
Qy 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCCCAACATGATGTTTCTGTGACCAAC 4800
Qy 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACTTGGGAGCTGGGGCCGGGAAAGACG 4860
Db 4801 CCTCTAACACCGTCCCTCTGGGAGCAACACGAGACTTGGGAGCTGGGGCCGGGAAAGACG 4860
Qy 4861 CCCGTTGATGATGACAGAGCAGCGGCATCATCAATGATCCGACTGCGATATGACACCC 4920
Db 4861 CCCGTTGATGATGACAGAGCAGCGGCATCATCAATGATCCGACTGCGATATGACACCC 4920
Qy 4921 AGCGTTGGCAGGCGCGCTGTTCTTAAGGCCCAACAGACTCTACTGCGGGCGGTGTTGG 4980
Db 4921 AGCGTTGGCAGGCGCGCTGTTCTTAAGGCCCAACAGACTCTACTGCGGGCGGTGTTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 3
CQ874961
LOCUS 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION CQ874961
VERSION CQ874961.1 GI:52748060
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Diamandis, E.P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
source I. 11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGSCCCAGAGTGAAGCAAGAGAGGAGTTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
Db 1 GGSCCCAGAGTGAAGCAAGAGAGGAGTTGAGAGTCCCTCTGCAAGTGGCTTGAGTC 60
Qy 61 TCCCTGCTCCCTAAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTGCTCCCTAAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGATAAACAACTACAGAAACACAGAGAGAAC 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGATAAACAACTACAGAAACACAGAGAGAAC 180
Qy 181 ACAGAGAGCTGGGACACAGGAGCACACAGAGTCAGAGAGAAAGAGAGATAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGAGCACACAGAGTCAGAGAGAAAGAGAGATAGAGAG 240
Qy 241 ACACAAATGGAGACACAGAGAGGTGTAAGAAAGAGAGATTAAACAGAGTCCACAGC 300
```



[illegible]

Dd	1321	AAAAGACAGAGAGAGAGAGACAGACAGAGAGAGAGAGAGAGCGGATAGANAAGA	1380
Qy	1381	GAGAGAGGGGTGGAGAGAGACACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA	1440
Dd	1381	GAGAGAGGGGTGGAGAGAGACACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA	1440
Qy	1441	GAACCACAGAGAGANTGGAAAGACTCTGAGAAAAAACCCAGAGACAAGNTGGAAGAGG	1500
Dd	1441	GAACCACAGAGAGANTGGAAAGACTCTGAGAAAAAACCCAGAGACAAGNTGGAAGAGG	1500
Qy	1501	AGTATCCAGGGTGAAACAGACAGTGGTGGAAATGAGCAAAATGCAGAGAAGAAACCAAGCAA	1560
Dd	1501	AGTATCCAGGGTGNAACAGACAGTGGTGGAAATGAGCAAAATGCAGAGAGNAAACCAAGCAA	1560
Qy	1561	TCCAGGCCCAAGAATAAGTCACCCAGAGTTGGTGAGAAGCCAGATCCTTAAGSCTCGGGG	1620
Dd	1561	TCCAGGCCCAAGAAATAGTCACCCAGAGTTGGTGAGAAGCCAGATCCTTAAGSCTCGGGG	1620
Qy	1621	AGGCAGGGAAGGGGCTGGCCCTGGCTTCCGAGACCCCCTCCCCATTCTCCGGSCCAGGGAG	1680
Dd	1621	AGGCAGGGAAGGGGCTGGCCCTGGCTTCCGAGACCCCCTCCCCATTCTCCGGSCCAGGGAG	1680
Qy	1681	GTAGGGAGTGACATTCCGGACTGGSTGGSGGTGCTCTGGSGGTGGAGATAGGGGGAGCA	1740
Dd	1681	GTAGGGAGTGACATTCCGGACTGGSTGGSGGTGCTCTGGSGGTGGAGATAGGGGGAGCA	1740
Qy	1741	GGAGGAGCTATTGCTAAGGCCCGCATAGAGCACCTCATTTGCCCGGGAATGTGCCCCAGGGAG	1800
Dd	1741	GGAGGAGCTATTGCTAAGGCCCGCATAGAGCACCTCATTTGCCCGGGAATGTGCCCCAGGGAG	1800
Qy	1801	CAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGAGAGGAGCGAGTGGCCAGGA	1860
Dd	1801	CAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGAGAGGAGCGAGTGGCCAGGA	1860
Qy	1861	AGGCACAGGCTCAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCACCCCCTACTCGG	1920
Dd	1861	AGGCACAGGCTCAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCACCCCCTACTCGG	1920
Qy	1921	GGGACAGGCGAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCGAGGAAGAGAGGTGTCT	1980
Dd	1921	GGGACAGGCGAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCGAGGAAGAGAGGTGTCT	1980
Qy	1981	GTGGGTCTTGACCCACATCTTTCTCTGTCCCTCTCTGTCCCTGTCTGAGGCTGTCTAGA	2040
Dd	1981	GTGGGTCTTGACCCACATCTTTCTCTGTCCCTCTCTGTCCCTGTCTGAGGCTGTCTAGA	2040
Qy	2041	CTCCATCTTCTGAAATTTCTATAGTGCCTGGGTCTCAGCGCAGTCCGATGTGGCCCGTC	2100
Dd	2041	CTCCATCTTCTGAAATTTCTATAGTGCCTGGGTCTCAGCGCAGTCCGATGTGGCCCGTC	2100
Qy	2101	CTTTGGTTCTCTCTACCTGGGGAAATAAGGTAGGGGAGGGGAAAGTGGGTTAAGG	2160
Dd	2101	CTTTGGTTCTCTCTACCTGGGGAAATAAGGTAGGGGAGGGGAAAGTGGGTTAAGG	2160
Qy	2161	GCTCCC CGGATTCGGCTTGGGCTCCCAACCTCTGACATTCGCCATTCAGGTGCAGCGGCC	2220
Dd	2161	GCTCCC CGGATTCGGCTTGGGCTCCCAACCTCTGACATTCGCCATTCAGGTGCAGCGGCC	2220
Qy	2221	ATGGCTACAGCNAGACCCCTCGATGGGTGCTCTGTGCTCTGTATCAGAGCTTGCCTT	2280
Dd	2221	ATGGCTACAGCNAGACCCCTCGATGGGTGCTCTGTGCTCTGTATCAGAGCTTGCCTT	2280
Qy	2281	CTGGGGGTTCACAGGTAAACAGAACTCTGGGGTGGGAGGGTTGTGGGATTCGGAGGACTGT	2340
Dd	2281	CTGGGGGTTCACAGGTAAACAGAACTCTGGGGTGGGAGGGTTGTGGGATTCGGAGGACTGT	2340
Qy	2341	CTCTGCGGCATCAGAGCGCTGTCCCTCGGGGAACTGTGTGAGCCTGGGCAATGACTCCGG	2400
Dd	2341	CTCTGCGGCATCAGAGCGCTGTCCCTCGGGGAACTGTGTGAGCCTGGGCAATGACTCCGG	2400
Qy	2401	GACCGGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTTGTGCCATCGTATGTGGCCCTGT	2460
Dd	2401	GACCGGGTGAATGTGAGTCTCTGTCTGTACTTGTGGTTGTGCCATCGTATGTGGCCCTGT	2460

QY 2461 GACTGCCACGGTGTGTGTCGGGAGGGGATGECCTTTTCCATATCAGGTGACTGTGCGG 2520  
Db 2461 GACTGCCACGGTGTGTGTCGGGAGGGGATGECCTTTTCCATATCAGGTGACTGTGCGG 2520  
QY 2521 CAGGTGGCAGTACCCCTTTGAGGCTGTGTGTGTTGTGATGTGTGTGCAATTAAG 2580  
Db 2521 CAGGTGGCAGTACCCCTTTGAGGCTGTGTGTGTTGTGATGTGTGTGCAATTAAG 2580  
QY 2581 ATTGTGTGGCTCCACAGCTGTGTGGGTGAATGATGTAGCACTGGGGGTGTTCACTGT 2640  
Db 2581 ATTGTGTGGCTCCACAGCTGTGTGGGTGAATGATGTAGCACTGGGGGTGTTCACTGT 2640  
QY 2641 GTGTTTGGCTGTGTGTGTGTGCTTTGGCAATTGATATGACTGCAGGTATCTGCAGTCTCTG 2700  
Db 2641 GTGTTTGGCTGTGTGTGTGTGCTTTGGCAATTGATATGACTGCAGGTATCTGCAGTCTCTG 2700  
QY 2701 TCCCTGAGGTCCTGGGATTTGGTGCACAAAGTGGTCAATCACCATGAAAGCTGTGACT 2760  
Db 2701 TCCCTGAGGTCCTGGGATTTGGTGCACAAAGTGGTCAATCACCATGAAAGCTGTGACT 2760  
QY 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGAGTGTATGATGATGATGAT 2820  
Db 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTTGGCTGAGTGTGAGTGTATGATGATGATGAT 2820  
QY 2821 TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880  
Db 2821 TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880  
QY 2881 TCTGTGTGAGCCCGTGTAAATGCTACTGTATGTGTGATGTGTGAGTGTGTGTGAGT 2940  
Db 2881 TCTGTGTGAGCCCGTGTAAATGCTACTGTATGTGTGATGTGTGAGTGTGTGTGAGT 2940  
QY 2941 TTCTGTCTCTGCTGGAGGATAGAGGTGCAGGGTAGCTATCTCTGGGAGTGGGTGC 3000  
Db 2941 TTCTGTCTCTGCTGGAGGATAGAGGTGCAGGGTAGCTATCTCTGGGAGTGGGTGC 3000  
QY 3001 CAGGTGACTGACTTTCAGTGTGTGCTGTGTGCAGAGAGATGTGGCAGTCTGAAATC 3060  
Db 3001 CAGGTGACTGACTTTCAGTGTGTGCTGTGTGCAGAGAGATGTGGCAGTCTGAAATC 3060  
QY 3061 TGTGCACACAGGCACTGTGCGTGGGCACTGAGACACTGTGGATGAGGGGTGCGATCCC 3120  
Db 3061 TGTGCACACAGGCACTGTGCGTGGGCACTGAGACACTGTGGATGAGGGGTGCGATCCC 3120  
QY 3121 GCTAGGCTGCCGGAGCGTGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180  
Db 3121 GCTAGGCTGCCGGAGCGTGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACCTGTG 3180  
QY 3181 GAGGCAACATGGGCGTGTCTGCAGAACTGGCTGCGTGTGGCTGTACTGTCTGTGTGC 3240  
Db 3181 GAGGCAACATGGGCGTGTCTGCAGAACTGGCTGCGTGTGGCTGTACTGTCTGTGTGC 3240  
QY 3241 GCGTGGTCTTGGGGTGAATTCGTAATGATGTGGTGCAGGGCCATCAGCAAGGGTAA 3300  
Db 3241 GCGTGGTCTTGGGGTGAATTCGTAATGATGTGGTGCAGGGCCATCAGCAAGGGTAA 3300  
QY 3301 GAACAGGCGGGCGGTGGCTCACCGCTGTATCCAGCCCTTTTGGGAGGCGGAGCA 3360  
Db 3301 GAACAGGCGGGCGGTGGCTCACCGCTGTATCCAGCCCTTTTGGGAGGCGGAGCA 3360  
QY 3361 GCGGATCAGCTGAGGTGGGAGATCGAGGCGAGCTGTGACCAATGGAGAACCCCGTCT 3420  
Db 3361 GCGGATCAGCTGAGGTGGGAGATCGAGGCGAGCTGTGACCAATGGAGAACCCCGTCT 3420  
QY 3421 CTACTAAAAATACAAAAAATTTAGTGTGTGTGGCGCGTCCCTGTAATCCAGCTACTC 3480  
Db 3421 CTACTAAAAATACAAAAAATTTAGTGTGTGTGGCGCGTCCCTGTAATCCAGCTACTC 3480  
QY 3481 GGGAGACTGGGCGAGAAAAATCGTTGAAACCCGGAGGTGGAGGTTCGGTGGAGCGGCA 3540  
Db 3481 GGGAGACTGGGCGAGAAAAATCGTTGAAACCCGGAGGTGGAGGTTCGGTGGAGCGGCA 3540

QY 3541 TCGCGCCATTGCATCCAGCTTGGCAACAGAGCGAAAACTCCGTCTTCGAAAGAAAAA 3600  
Db 3541 TCGCGCCATTGCATCCAGCTTGGCAACAGAGCGAAAACTCCGTCTTCGAAAGAAAAA 3600  
QY 3601 GAAAAAAGGTTAAGAACAGTGAATGGGCAACGGAGGACTGATGATGGAGTGGGG 3660  
Db 3601 GAAAAAAGGTTAAGAACAGTGAATGGGCAACGGAGGACTGATGATGGAGTGGGG 3660  
QY 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAGGCAT 3720  
Db 3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAGGCAT 3720  
QY 3721 GTTTTCATCTGAGAAATTCAGAAACCTAGGCTCTCTTCCCTCCATGTGGCCCCCTAAG 3780  
Db 3721 GTTTTCATCTGAGAAATTCAGAAACCTAGGCTCTCTTCCCTCCATGTGGCCCCCTAAG 3780  
QY 3781 CTGAGCCCTTCTTCTCTGTGCTCTTTCGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840  
Db 3781 CTGAGCCCTTCTTCTCTGTGCTCTTTCGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840  
QY 3841 CCACCTCTTCTCAACACGCCCCCTTAGGCCAGACTCTAGTGGACCCCGCTAAGGCCA 3900  
Db 3841 CCACCTCTTCTCAACACGCCCCCTTAGGCCAGACTCTAGTGGACCCCGCTAAGGCCA 3900  
QY 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGGTACCTTCTAGAAACCCCTTCAA 3960  
Db 3901 CACCCCTTTGGGCGAGGCTCCACCCCTATTCTGTGGGTACCTTCTAGAAACCCCTTCAA 3960  
QY 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGCTTGTCTCTCTCCAGGCTGGAG 4020  
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGCTTGTCTCTCTCCAGGCTGGAG 4020  
QY 4021 TGCAGTGGGCTGATCTCGGCTCACTGCACCTCTGCTCCAGGTTCAAGTGAATCTCGT 4080  
Db 4021 TGCAGTGGGCTGATCTCGGCTCACTGCACCTCTGCTCCAGGTTCAAGTGAATCTCGT 4080  
QY 4081 GCCTCCACCTCTCTGAGTAGCTGGGATTAACAGGTGCGCGCCACACACGCTGGCTTAATTTT 4140  
Db 4081 GCCTCCACCTCTCTGAGTAGCTGGGATTAACAGGTGCGCGCCACACACGCTGGCTTAATTTT 4140  
QY 4141 GTGCTTTTGTAGTAGACAGGGTTTCACTTTGTTGGCAGGCTGGTCTCAAACCTCCAAAC 4200  
Db 4141 GTGCTTTTGTAGTAGACAGGGTTTCACTTTGTTGGCAGGCTGGTCTCAAACCTCCAAAC 4200  
QY 4201 TCAAGTGTACCGCCACCTCGGCTCCAGAGTGTCTGGGGTTACAGGCTGAGCGCACCGC 4260  
Db 4201 TCAAGTGTACCGCCACCTCGGCTCCAGAGTGTCTGGGGTTACAGGCTGAGCGCACCGC 4260  
QY 4261 CCCAGCCCCAAGTCAGAGCTCTTTATAGGAGACTCTAACTGTAACCTGACCTGGGCC 4320  
Db 4261 CCCAGCCCCAAGTCAGAGCTCTTTATAGGAGACTCTAACTGTAACCTGACCTGGGCC 4320  
QY 4321 CTAACTAAGTCAATTCGAAACCCCTTCTGCTCCAGGCTGACCCCACTCACTGAGGCC 4380  
Db 4321 CTAACTAAGTCAATTCGAAACCCCTTCTGCTCCAGGCTGACCCCACTCACTGAGGCC 4380  
QY 4381 TGACCCCACTCTTGTGAGACAGTTCCATCCCTAAAGCCCTGCTGCTCCGATCCCGCAG 4440  
Db 4381 TGACCCCACTCTTGTGAGACAGTTCCATCCCTAAAGCCCTGCTGCTCCGATCCCGCAG 4440  
QY 4441 GCTCCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAACTCTGTACCCAAT 4500  
Db 4441 GCTCCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAACTCTGTACCCAAT 4500  
QY 4501 TTTTACCCCTCAATGTAGTTCTAGCCAAATTCAGGAAATCTGTGAGGTCAGTTAGTCCA 4560  
Db 4501 TTTTACCCCTCAATGTAGTTCTAGCCAAATTCAGGAAATCTGTGAGGTCAGTTAGTCCA 4560  
QY 4561 GTAAACCCCTACCTGAGGCTCTGCTCTTGTAGCTTGGAGCTGGGGTTGAGAGGTGCCA 4620  
Db 4561 GTAAACCCCTACCTGAGGCTCTGCTCTTGTAGCTTGGAGCTGGGGTTGAGAGGTGCCA 4620  
QY 4621 CTCCTTATTTCCAGGCGCTTCCCTTCCCTTCCCTCAGCATGTGACAGACACCCACCTCTAGCT 4680



```
Db 4621 CTCCTATTCTCCAGGCTGCCCCCTGCCCCCTGAGCATGTGACACACCCACCTCTAGCT 4680
Qy 4681 GGTCTGGCTCTTGTAGTCTGAACCCACCCAGCCCAAGCCCGCTCTCTAGCCCGGCC 4740
Db 4681 GGTCTGGCTCTTGTAGTCTGAACCCACCCAGCCCAAGCCCGCTCTCTAGCCCGGCC 4740
Qy 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTGCGCCCAACATGATGTTTCTGTGACACC 4800
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTGCGCCCAACATGATGTTTCTGTGACACC 4800
Qy 4801 CCTCTAACACCGTCTGCTCTGGAGCAACACAGGACCTGGAGCTGGGCGCGGGAAGACG 4860
Db 4801 CCTCTAACACCGTCTGCTCTGGAGCAACACAGGACCTGGAGCTGGGCGCGGGAAGACG 4860
Qy 4861 CCCGCTGGATGACAGAGCAGCGCCGATCATCAATGATGATGATGATGATGATGATGATG 4920
Db 4861 CCCGCTGGATGACAGAGCAGCGCCGATCATCAATGATGATGATGATGATGATGATGATG 4920
Qy 4921 AGCGTGGAGCGCGCTGTTGCTTAAGGCCCAACACGCTCTACTGCGGGCGGTGTGG 4980
Db 4921 AGCGTGGAGCGCGCTGTTGCTTAAGGCCCAACACGCTCTACTGCGGGCGGTGTGG 4980
Qy 4981 TGCATCCACAGTGGCTGCTC 5000
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 4
AF135028
LOCUS
DEFINITION Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 11570)
Yousef, G.M. and Diamandis, E.P.
The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
J. Biol. Chem. 274 (53), 37511-37516 (1999)
20076408
PUBMED
10608802
REFERENCE
2 (bases 1 to 11570)
Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obierzu, C.V.
The new human kallikrein gene family: implications in
carcinogenesis
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
21121728
PUBMED
10675891
REFERENCE
3 (bases 1 to 11570)
Yousef, G.M., Luo, L.Y. and Diamandis, E.P.
Direct Submission
Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
Location/Qualifiers
1..11570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101..2331,2210..2293,4762..5023,5763..6019,
6105..6238,11092..11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221..2293,4762..5023,5763..6019,6105..6238,
11092..11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
```

```
/protein_id="AAD26429.1"
/db_xref="GI:4589283"
/translation="MATAPRPMWVLCALITALLGLVTEHVLANNVSDCHPSNTVPS
GSDQDLGAGGAGEDARSDSSRIINGSDDMHQTPWQALLRLPNQYCGAVLHPQW
LLTAHKKRVFRVRLHYSLSFVYSGQQMFQGVKSIPIHPGYSHPGSHNDMLIKLN
LRIRPTKDVQPIINVSHCPAGTKCLVSGWGTTSKPOVHPFKVLOCLNLSVLSOKRCE
DAYPROIDDITMFCAGDKAGRDSCOGDSGSPVVCNGSLQGLVSGDYPFCARPNRGVVT
NLCKFYKWIQETIQANS"

ORIGIN
Query Match 100.0%; Score 5000; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCCCAAGTGAAGGCAAGAGAGAGTTCCTCTGCAAGTGGCTTGTAGTC 60
Db 1 GGGCCCAAGTGAAGGCAAGAGAGAGTTCCTCTGCAAGTGGCTTGTAGTC 60
Qy 61 TCCCTCCCTAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTCCCTAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAGAGAGAGAGAGAG 240
Qy 241 ACACAAATGAGAGACAGAGAGTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 ACACAAATGAGAGACAGAGAGTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Qy 301 AAAGGGGACAGAGACAGTCTTTCAGGGTGGTGTCTATGATCATCTCTCTTTT 360
Db 301 AAAGGGGACAGAGACAGTCTTTCAGGGTGGTGTCTATGATCATCTCTCTTTT 360
Qy 361 TTTTCTTTTCTTTTGTAGACGGAGTCTCGCTCTGTCTGCCAGGTCGAGTCAGTCG 420
Db 361 TTTTCTTTTCTTTTGTAGACGGAGTCTCGCTCTGTCTGCCAGGTCGAGTCAGTCG 420
Qy 421 GATCTCGGCTCAGTCGAGCTCCGCTCCGGGTTCAGGCAATCTCTCTGCTCAGCTC 480
Db 421 GATCTCGGCTCAGTCGAGCTCCGCTCCGGGTTCAGGCAATCTCTCTGCTCAGCTC 480
Qy 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCCCGGCTAATTTTGTATTTTA 540
Db 481 CCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCCCGGCTAATTTTGTATTTTA 540
Qy 541 GTAGAGACGGGGTTTCAACCGTTTTCAGCGGGATGGCTTCGATCTCTGACCTCGTATCC 600
Db 541 GTAGAGACGGGGTTTCAACCGTTTTCAGCGGGATGGCTTCGATCTCTGACCTCGTATCC 600
Qy 601 CCCCGCTCCGCTCCCAAGTGTGGGATTCAGGCGTCAGGCGCCCGCCCGCCCATG 660
Db 601 CCCCGCTCCGCTCCCAAGTGTGGGATTCAGGCGTCAGGCGCCCGCCCGCCCATG 660
Qy 661 ATCATCTTCTGACTGATGTCGACAGTACTTAAAGCCATCAGACTCTACCTTTTA 720
Db 661 ATCATCTTCTGACTGATGTCGACAGTACTTAAAGCCATCAGACTCTACCTTTTA 720
Qy 721 AATATGAGTTTGGGCCAGGACCGTGGCTCATGCTGTAAATTCAGACTTTGGGAGGC 780
Db 721 AATATGAGTTTGGGCCAGGACCGTGGCTCATGCTGTAAATTCAGACTTTGGGAGGC 780
Qy 781 AGAGTGGGTGAATCACTTTCAGGCGAGAGTTTTCAGACCCAGCTTTCAGGAGGC 840
Db 781 AGAGTGGGTGAATCACTTTCAGGCGAGAGTTTTCAGACCCAGCTTTCAGGAGGC 840
Qy 841 CTCTGCTCTTACTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 CTCTGCTCTTACTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
```

Qy	901	ACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGCG	960
Dd	901	ACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAAACCTCGAGGCG	960
Qy	961	GAGGTTCAGTGGCGCGAGATCAATCACCGCCCTCCAGCCCTGGCGACAGCAAGACT	1020
Dd	961	GAGGTTCAGTGGCGCGAGATCAATCACCGCCCTCCAGCCCTGGCGACAGCAAGACT	1020
Qy	1021	CTGTCTCAAAATAAATAAACAACGAAACGAAACAGCTTTGTCTGACCTAGTTATATCT	1080
Dd	1021	CTGTCTCAAAATAAATAAACAACGAAACAGCTTTGTCTGACCTAGTTATATCT	1080
Qy	1081	AAAAAATAAATGCTGTCAACAAATAGACGAAAGTGAATAAAGGAAATAAATGGGCCA	1140
Dd	1081	AAAAAATAAATGCTGTCAACAAATAGACGAAAGTGAATAAAGGAAATAAATGGGCCA	1140
Qy	1141	AGAACTCTAAGGTATATTTGACAAATCTTTCAGAACCTTTTAAAAAGAAAGATCAAGA	1200
Dd	1141	AGAACTCTAAGGTATATTTGACAAATCTTTCAGAACCTTTTAAAAAGAAAGATCAAGA	1200
Qy	1201	GGCATAGAAACACAGGAGGAGAACAGGAGACAGAAACACCTGTGCGCCCAAGGAAACAA	1260
Dd	1201	GGCATAGAAACACAGGAGGAGAACAGGAGACAGAAACACCTGTGCGCCCAAGGAAACAA	1260
Qy	1261	ACAAAGGCTCTTAAGACACAGAGGAGGAGACAGAGAGAGTGAGTGAGAGACAGACAGAG	1320
Dd	1261	ACAAAGGCTCTTAAGACACAGAGGAGGAGACAGAGAGAGTGAGTGAGAGACAGACAGAG	1320
Qy	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGAGATAGAAGA	1380
Dd	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAGGAGATAGAAGA	1380
Qy	1381	GAGAGAGGCTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
Dd	1381	GAGAGAGGCTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGA	1440
Qy	1441	GAAACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACCCAGAGACAAAGATGGAAGAAG	1500
Dd	1441	GAAACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACCCAGAGACAAAGATGGAAGAAG	1500
Qy	1501	AGTATCGAGGGTGAAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
Dd	1501	AGTATCGAGGGTGAAACAGACAGTGGTGAATGAGCAAAATGACAGAGAAAGCAAGCAA	1560
Qy	1561	TCCAGGCCCAAGATAGTGACCCAGAGTGGTGAGAGCCAGATCCCTTAAGGCTGGGG	1620
Dd	1561	TCCAGGCCCAAGATAGTGACCCAGAGTGGTGAGAGCCAGATCCCTTAAGGCTGGGG	1620
Qy	1621	AGGCAGGGAAGGGCTGGCCCTGGCTTCCGAGAGACCCCTCCCAATCTCCGGGCCAGGGAG	1680
Dd	1621	AGGCAGGGAAGGGCTGGCCCTGGCTTCCGAGAGACCCCTCCCAATCTCCGGGCCAGGGAG	1680
Qy	1681	GTAAGGAGTGACATTCGGAATCGGGTGGGGTGCTCTGGGGTGGAGATAGGGGGAGCA	1740
Dd	1681	GTAAGGAGTGACATTCGGAATCGGGTGGGGTGCTCTGGGGTGGAGATAGGGGGAGCA	1740
Qy	1741	GGAGGAGCTATTGCTAAGGCCCGATAGCACTCATTTGCCCGGAAATGTGCCCCAGGGAG	1800
Dd	1741	GGAGGAGCTATTGCTAAGGCCCGATAGCACTCATTTGCCCGGAAATGTGCCCCAGGGAG	1800
Qy	1801	CAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGGAGGAGCGAGTGCCCAGGA	1860
Dd	1801	CAGTGGGTGGTTATAACTCAGGCCCGGTGCCAGAGCCAGGAGGAGCGAGTGCCCAGGA	1860
Qy	1861	AGGCAGGCTGAGAAAGTCTGCGGCTGAGCTGGAGCAATATCCCAACCCCTACCTGG	1920
Dd	1861	AGGCAGGCTGAGAAAGTCTGCGGCTGAGCTGGAGCAATATCCCAACCCCTACCTGG	1920
Qy	1921	GGGACAGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGACGAGAGAGGTTCT	1980
Dd	1921	GGGACAGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGACGAGAGAGGTTCT	1980

Qy	1981	GTGCGTCTGCGACCCACATCTTTCTGTGTCCTCTGTGCCCTCTGTGAGGCTGCTAGA	2040
Dd	1981	GTGCGTCTGCGACCCACATCTTTCTGTGTCCTCTGTGCCCTCTGTGAGGCTGCTAGA	2040
Qy	2041	CTCTATCTTCTGAAATCTATAGTGCTGGGTCTAGCGCAGTGCCGATGGTGCCCGTCT	2100
Dd	2041	CTCTATCTTCTGAAATCTATAGTGCTGGGTCTAGCGCAGTGCCGATGGTGCCCGTCT	2100
Qy	2101	CTTGTGTTCTCTCTACCTGGGGAATAAGGTAGGGAGGGAGGGGAAGTGGGTTAAGG	2160
Dd	2101	CTTGTGTTCTCTCTACCTGGGGAATAAGGTAGGGAGGGAGGGGAAGTGGGTTAAGG	2160
Qy	2161	GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGCGAGCGGCC	2220
Dd	2161	GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTTCCCATCCAGGTGCGAGCGGCC	2220
Qy	2221	ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTCTGATCAAGCTTGCTT	2280
Dd	2221	ATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTCTGATCAAGCTTGCTT	2280
Qy	2281	CTGGGGTCAACAGTAACAGAACTCTGGGGTGGAGGGTGTGGGATTCGGAGGACTGT	2340
Dd	2281	CTGGGGTCAACAGTAACAGAACTCTGGGGTGGAGGGTGTGGGATTCGGAGGACTGT	2340
Qy	2341	CTCTGGGGCACTAGAGCGCTGCCCTGGGGAACTGTGTGAGCCTGGGCATGACTCCGG	2400
Dd	2341	CTCTGGGGCACTAGAGCGCTGCCCTGGGGAACTGTGTGAGCCTGGGCATGACTCCGG	2400
Qy	2401	GACGGGTGAATGTGAGTCTCTGTCTGTGTAATGTGGTGTGGATTCGATGTGGCCCTGT	2460
Dd	2401	GACGGGTGAATGTGAGTCTCTGTCTGTGTAATGTGGTGTGGATTCGATGTGGCCCTGT	2460
Qy	2461	GACTGCCACGCTGTGTGGGGAGGGGATGCTTTCCCATCATCAGGTGACTGTGCGG	2520
Dd	2461	GACTGCCACGCTGTGTGGGGAGGGGATGCTTTCCCATCATCAGGTGACTGTGCGG	2520
Qy	2521	CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGGTGTGGTGTGGATTTAAG	2580
Dd	2521	CAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGGTGTGGTGTGGATTTAAG	2580
Qy	2581	ATTGTGTGGCTCCACAGCTGTGGGTGAATGCAATGTAAGCACTGGGGGTGTTCACTGT	2640
Dd	2581	ATTGTGTGGCTCCACAGCTGTGGGTGAATGCAATGTAAGCACTGGGGGTGTTCACTGT	2640
Qy	2641	GTCTTTGGCTGTGTGGTGTGACCTTGCAATGATATGACATGCAAGTATCTGCAAGTCTG	2700
Dd	2641	GTCTTTGGCTGTGTGGTGTGACCTTGCAATGATATGACATGCAAGTATCTGCAAGTCTG	2700
Qy	2701	TCCCTGAGGTCCCGGATTCGCTGCAACAAAAGTGGTCAATCACCATGGAAGCTGTGACT	2760
Dd	2701	TCCCTGAGGTCCCGGATTCGCTGCAACAAAAGTGGTCAATCACCATGGAAGCTGTGACT	2760
Qy	2761	GTGTGCTTGTGAGGGGATTAATGTAATGTGGCTGAGTGTGAGCTTATGGATGCCGTA	2820
Dd	2761	GTGTGCTTGTGAGGGGATTAATGTAATGTGGCTGAGTGTGAGCTTATGGATGCCGTA	2820
Qy	2821	TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGCACTGTGTGTG	2880
Dd	2821	TTTGTGACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGCACTGTGTGTG	2880
Qy	2881	TCCTGTGAGGGCTGTAAATGCTACTGTATGTGTGATGGTGTGAGCTGTGTCTGGAGT	2940
Dd	2881	TCCTGTGAGGGCTGTAAATGCTACTGTATGTGTGATGGTGTGAGCTGTGTCTGGAGT	2940
Qy	2941	TTCTGTCTGTGCTGAGGGGATAGAGGTGTCAGGGGTAGCTATCTCTGGGAGATGGGTGC	3000
Dd	2941	TTCTGTCTGTGCTGAGGGGATAGAGGTGTCAGGGGTAGCTATCTCTGGGAGATGGGTGC	3000
Qy	3001	CAGGTGACTGACTTGCAGTGTGCTGTGTGCAAGAGTATGTGCACTCTGCAACATC	3060
Dd	3001	CAGGTGACTGACTTGCAGTGTGCTGTGTGCAAGAGTATGTGCACTCTGCAACATC	3060
Qy	3061	TGTGCAACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC	3120

Db 3061 TGTCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGATGAGGGTGTGCGATCCC 3120  
Qy 3121 GCTAGGCTGCCGGAGGGTGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180  
Db 3121 GCTAGGCTGCCGGAGGGTGTGTACCTGGAGACAGAGCTGTATGTTAGCTGCACTGTGTG 3180  
Qy 3181 GAGGCAACATGGGCGTGTCTCAGAACTGCGTGGTGTCTGGCTGTACTCTGTGTGTC 3240  
Db 3181 GAGGCAACATGGGCGTGTCTCAGAACTGCGTGGTGTCTGGCTGTACTCTGTGTGTC 3240  
Qy 3241 GCGTGGTCTTTGGGGTGAGTTTCGTAATGATGTGGTGGCCAGGGCCATCAGCAAGGGTAA 3300  
Db 3241 GCGTGGTCTTTGGGGTGAGTTTCGTAATGATGTGGTGGCCAGGGCCATCAGCAAGGGTAA 3300  
Qy 3301 GAAACAGGCGGGCGGGTGGCTCAGCGCTGTATCCAGGCCCTTTGGGAGGCCGAGGCA 3360  
Db 3301 GAAACAGGCGGGCGGGTGGCTCAGCGCTGTATCCAGGCCCTTTGGGAGGCCGAGGCA 3360  
Qy 3361 GCGGATCACCTGAGGTCGGGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420  
Db 3361 GCGGATCACCTGAGGTCGGGAGATCGAGGCCAGCTGACCAACATGGAGAACCCCGTCT 3420  
Qy 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGGTGGCGGTGCTGTATCCAGCTACTC 3480  
Db 3421 CTACTAAAAATACAAAAATTTAGCTGTGTGGTGGCGGTGCTGTATCCAGCTACTC 3480  
Qy 3481 GGGAGACTGGGGCAGAAAAATTCGCTTGAAACCCGGGAGGTGAGGTTGGGTGAGCCGAGA 3540  
Db 3481 GGGAGACTGGGGCAGAAAAATTCGCTTGAAACCCGGGAGGTGAGGTTGGGTGAGCCGAGA 3540  
Qy 3541 TCGCGCATTCGACTCCAGCTCGGCAACAGAGCGAACTCCGCTCGAAGAAAAAAA 3600  
Db 3541 TCGCGCATTCGACTCCAGCTCGGCAACAGAGCGAACTCCGCTCGAAGAAAAAAA 3600  
Qy 3601 GAAAAAAGGGTGAAGAACAGTGAATGGGCAAGGAGGACTGATGAGTGGGGC 3660  
Db 3601 GAAAAAAGGGTGAAGAACAGTGAATGGGCAAGGAGGACTGATGAGTGGGGC 3660  
Qy 3661 ATGATGATGATCTGTAGTCTGTGTGTGAGAGGAGAGATTGACAGATTGAGAAGGCAT 3720  
Db 3661 ATGATGATGATCTGTAGTCTGTGTGTGAGAGGAGAGATTGACAGATTGAGAAGGCAT 3720  
Qy 3721 GTTTTCATCATGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGCCCTTAAG 3780  
Db 3721 GTTTTCATCATGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGGCCCTTAAG 3780  
Qy 3781 CTGAGCCCTCTTTCTGCTGCTCTTTCGGAAACCTAGCTCCGCCATGAGCTCTGACC 3840  
Db 3781 CTGAGCCCTCTTTCTGCTGCTCTTTCGGAAACCTAGCTCCGCCATGAGCTCTGACC 3840  
Qy 3841 CCACCTCTTTCTCAACACGCCCTTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900  
Db 3841 CCACCTCTTTCTCAACACGCCCTTAGGCCAGACTCTAGTGGACCCCGCTTAAGGCCA 3900  
Qy 3901 CACCCCTTTGGCCAGGCTCAGCCCTTATCTGTGGTACTCTTGTAGAACCCCTTCAA 3960  
Db 3901 CACCCCTTTGGCCAGGCTCAGCCCTTATCTGTGGTACTCTTGTAGAACCCCTTCAA 3960  
Qy 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGCTGGAG 4020  
Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGTCTCTCTCTCCAGGCTGGAG 4020  
Qy 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGCT 4080  
Db 4021 TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGCT 4080  
Qy 4081 GCCTCCACTCTCAGTGTGAGTGGGATACAGGTGGCGGCCACACCGCTGGCTAATTTTT 4140  
Db 4081 GCCTCCACTCTCAGTGTGAGTGGGATACAGGTGGCGGCCACACCGCTGGCTAATTTTT 4140  
Qy 4141 GTGTCTTTAGTAGACAGGGTTTCACCTTTGTGGCAGGCTGTCTCAAACTCCCAACC 4200

Db 4141 GTGTCTTTAGTAGACAGGGTTTACCTTTGTTGGCCAGGCTGTCTCTCAAACTCCCAACC 4200  
Qy 4201 TCAGGTGATCGGCCCACTCGGCTCCAGAGTGTCTGGGTTACAGGCGTGAGCACCAGC 4260  
Db 4201 TCAGGTGATCGGCCCACTCGGCTCCAGAGTGTCTGGGTTACAGGCGTGAGCACCAGC 4260  
Qy 4261 CCCAGGCCCAAGTTCAGAGTCTTTATAGGAGACTCTAAACATGTAAACCTTGACCTGGCC 4320  
Db 4261 CCCAGGCCCAAGTTCAGAGTCTTTATAGGAGACTCTAAACATGTAAACCTTGACCTGGCC 4320  
Qy 4321 CTAACTAAGTCAATTTCAAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380  
Db 4321 CTAACTAAGTCAATTTCAAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380  
Qy 4381 TGACCCCACTCTTGAGACAGTTCATCCCTTAAGGCCCTGCTCTCCCTCCCACTCCAG 4440  
Db 4381 TGACCCCACTCTTGAGACAGTTCATCCCTTAAGGCCCTGCTCTCCCTCCCACTCCAG 4440  
Qy 4441 GCTCCAGCCCACTGAGCTTTGGGCACTACCTCTGAGCTTGTCCAGGAATCTCTGTACCCAAT 4500  
Db 4441 GCTCCAGCCCACTGAGCTTTGGGCACTACCTCTGAGCTTGTCCAGGAATCTCTGTACCCAAT 4500  
Qy 4501 TTATACCTCACTGTAGTCTTAGCCAAATTCAGGAATCTGTGAGGTTCCAGTTAGAGTCCA 4560  
Db 4501 TTATACCTCACTGTAGTCTTAGCCAAATTCAGGAATCTGTGAGGTTCCAGTTAGAGTCCA 4560  
Qy 4561 GTAAACCTACTGAGCCCTGGCTCTGTCTTGGAGCTTGGAGCTTGGAGGTTGAGGTTGAGG 4620  
Db 4561 GTAAACCTACTGAGCCCTGGCTCTGTCTTGGAGCTTGGAGCTTGGAGGTTGAGGTTGAGG 4620  
Qy 4621 CTCTTATTTCCAGGCGCTGCCCTGCCCTCAGCATGTCTAGACACCCACCTCTAGCT 4680  
Db 4621 CTCTTATTTCCAGGCGCTGCCCTGCCCTCAGCATGTCTAGACACCCACCTCTAGCT 4680  
Qy 4681 GGTCGSCCTCTTGTAGTCTGAAACCCCAAGCCCAAGCCCTCTGAGGCCCGGCC 4740  
Db 4681 GGTCGSCCTCTTGTAGTCTGAAACCCCAAGCCCAAGCCCTCTGAGGCCCGGCC 4740  
Qy 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCAACAAATGATGTTTCTGTGACCAAC 4800  
Db 4741 CAACCCATTTTCCGTTCCAGAGCATGTTCTCGCAACAAATGATGTTTCTGTGACCAAC 4800  
Qy 4801 CCTCTAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGGAAGACG 4860  
Db 4801 CCTCTAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGGAAGACG 4860  
Qy 4861 CCGGTCGGATGACAGCAGCAGCGCATCATCAATGATCCGACTGCGATATGCACACCC 4920  
Db 4861 CCGGTCGGATGACAGCAGCAGCGCATCATCAATGATCCGACTGCGATATGCACACCC 4920  
Qy 4921 AGCGTGGCAGGCGCGCTGTGCTAAGGCCCAACCAAGCTCTACTGCGGGCGGTGTGG 4980  
Db 4921 AGCGTGGCAGGCGCGCTGTGCTAAGGCCCAACCAAGCTCTACTGCGGGCGGTGTGG 4980  
Qy 4981 TGCATCCACAGTGGCTGCTC 5000  
Db 4981 TGCATCCACAGTGGCTGCTC 5000

## RESULT 5

AC011483/c

LOCUS

DEFINITION

AC011483

ACCESSION

AC011483.7

VERSION

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

AC011483 107487 bp DNA linear PRI 01-JUL-2002

Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.

AC011483

AC011483.7 GI:21637461

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 107487)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission





Qy 3840 CCCACCTCTTCTTCTCAACACACGCGCCCTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCC 3899  
Db 56380 CCCACCTCTTCTTCTCAACACACGCGCCCTAGGCCAGACTCTAGTGGACCCCGCCTTAAGGCC 56321  
Qy 3900 ACACCCCTTTGGGCGCAGCTCCACCCCTATTCTGTGGGTACCTTCTAGAACCCCTTCA 3959  
Db 56320 ACACCCCTTTGGGCGCAGCTCCACCCCTATTCTGTGGGTACCTTCTAGAACCCCTTCA 56261  
Qy 3960 AAGTCAGAGCTTTTCTTTTCTTTTGGAGACAGTCTCTCTCTCCAGGCTGGA 4019  
Db 56260 AAGTCAGAGCTTTTCTTTTCTTTTGGAGACAGTCTCTCTCTCCAGGCTGGA 56201  
Qy 4020 GTGCAGTGGGTGATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGATTTCTCG 4079  
Db 56200 GTGCAGTGGGTGATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGATTTCTCG 56141  
Qy 4080 TGCCTCCACCTCTCTAGTAGCTGGGATTAAGTGGCGGCGCACACCGCTGGCTAAATTT 4139  
Db 56140 TGCCTCCACCTCTCTAGTAGCTGGGATTAAGTGGCGGCGCACACCGCTGGCTAAATTT 56081  
Qy 4140 TGTGCTTTAGTACAGACAGGTTTCACTTGTGGCCAGGCTGCTCAAACTCCCAAC 4199  
Db 56080 TGTGCTTTAGTACAGACAGGTTTCACTTGTGGCCAGGCTGCTCAAACTCCCAAC 56021  
Qy 4200 CTCAAGTGATCCGCGCCACCTCGGCTCCAGAGTGTGGGTTACAGGCGTGAGCCACCG 4259  
Db 56020 CTCAAGTGATCCGCGCCACCTCGGCTCCAGAGTGTGGGTTACAGGCGTGAGCCACCG 55961  
Qy 4260 CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAAATGATTAACCTGACCTGGC 4319  
Db 55960 CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAAATGATTAACCTGACCTGGC 55901  
Qy 4320 CCTAACTAAGTCANTCCAAACCCCTTCTGCTCCAGGCTGACCCCACTCACTGAGGC 4379  
Db 55900 CCTAACTAAGTCANTCCAAACCCCTTCTGCTCCAGGCTGACCCCACTCACTGAGGC 55841  
Qy 4380 CTGACCCCACTTCTGTAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCA 4439  
Db 55840 CTGACCCCACTTCTGTAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCA 55781  
Qy 4440 GGCTCCAGCCCAAGCTTTGGCACTACCCCTGAGTGTTCAGAGGAATCTGTACCCAA 4499  
Db 55780 GGCTCCAGCCCAAGCTTTGGCACTACCCCTGAGTGTTCAGAGGAATCTGTACCCAA 55721  
Qy 4500 TTTTACCTCAGATGTAGTCTACGCAATCCAGGAATCTGTGAGGTCAGTTAGATCC 4559  
Db 55720 TTTTACCTCAGATGTAGTCTACGCAATCCAGGAATCTGTGAGGTCAGTTAGATCC 55661  
Qy 4560 AGTAACCTCAGCTGGGCTCTCTCTTGTAGCTTGTAGCCCTGGGCTTGTAGAGGTGCC 4619  
Db 55660 AGTAACCTCAGCTGGGCTCTCTCTTGTAGCTTGTAGCCCTGGGCTTGTAGAGGTGCC 55601  
Qy 4620 ACTCTTATCTCAGGCGCCCTGCCCTCGCCCTCAGCATGTCTAGACACCCACCCCTCTAGC 4679  
Db 55600 ACTCTTATCTCAGGCGCCCTGCCCTCGCCCTCAGCATGTCTAGACACCCACCCCTCTAGC 55541  
Qy 4680 TGGTCTGGCTCTTGTAGTCTGAACCCACCCCGAGCCAGCCCGGCTCTGAGCCCGCC 4739  
Db 55540 TGGTCTGGCTCTTGTAGTCTGAACCCACCCCGAGCCAGCCCGGCTCTGAGCCCGCC 55481  
Qy 4740 CCAACCATTTTCCGTTCCAGAGCATGTTCTGCCCAACATGATGTTCTCTGTGACCCAC 4799  
Db 55480 CCAACCATTTTCCGTTCCAGAGCATGTTCTGCCCAACATGATGTTCTCTGTGACCCAC 55421  
Qy 4800 CCTCTAACCCCTGCTCTGGGAGCAACAGGACCTGGAGCTGGGGCGGGGAGAGAC 4859  
Db 55420 CCTCTAACCCCTGCTCTGGGAGCAACAGGACCTGGAGCTGGGGCGGGGAGAGAC 55361  
Qy 4860 GCCCGTTCGATGACAGCAGGACCGCATCATCATGATCCGATCTCGATATGCACAC 4919  
Db 55360 GCCCGTTCGATGACAGCAGGACCGCATCATCATGATCCGATCTCGATATGCACAC 55301

Qy 4920 CAGCCGCTGGCAGCCGCGCTGTTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTG 4979  
Db 55300 CAGCCGCTGGCAGCCGCGCTGTTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTG 55241  
Qy 4980 GTGCATCCACAGTGGCTGCTC 5000  
Db 55240 GTGCATCCACAGTGGCTGCTC 55220  
RESULT 6  
AP243527/c  
LOCUS Homo sapiens serine protease gene cluster, complete sequence. 230000 bp DNA linear PRI 21-NOV-2000  
DEFINITION AP243527  
ACCESSION AP243527  
VERSION AP243527.1 GI:11244757  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 230000)  
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J.,  
Moss,P., Paepfer,B. and Wang,K.  
TITLE Sequencing and expression analysis of the serine protease gene  
cluster located in chromosome 19q13 region  
JOURNAL Gene 257 (1), 119-130 (2000)  
MEDLINE 20510030  
PUBMED 11054574  
REFERENCE 2 (bases 1 to 230000)  
AUTHORS Gan,L., Lee,I., Smith,R., Argonza-Barrett,R., Lei,H., McCuaig,J.,  
Moss,P., Paepfer,B. and Wang,K.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,  
Bothell, WA 98021, USA  
FEATURES  
Location/Qualifiers  
1..230000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13"  
complement(<6963..>11517)  
/genes="KLK1"  
9482..9641,11472..>11517)  
/genes="KLK1"  
/product="renal kallikrein"  
complement(Join(6963..7118,7668..7804,7923..8212,  
9482..9641,11472..11517))  
/genes="KLK1"  
/note="serine protease"  
/codon\_start=1  
/product="renal kallikrein"  
/protein\_id="AAG3353.1"  
/db\_xref="GI:11244758"  
/translation="MWFLVLCALSLGSGTGAAPQSIQRIIVGWGECGHSOPWQAALVH  
KFTPCGGILVHRQWVLTAACTSDNYQLWGRHNPDDNTAQFVHVSFPHPGEN  
MSLENNHTQADEYDSDHLMRLKTEPADITDAVKVELFTPEPEVGSFCLASGWS  
IPENFSFDLDQCVDLKLIPNDEKKAHVQKVTDFMLCVGHLEGGKDTCCVDSGGSL  
MCDGLVQVTSWGYVPCGTPKPSVAVRVLSYVKWIEDTIAENS"  
complement(Join(<13552..13704,14377..14530,14678..14917,  
15416..15569,19204..>19246))  
/product="ACO protease"  
complement(Join(13552..13704,14377..14530,14678..14917,  
15416..15569,19204..>19246))  
/notes="serine protease"  
/codon\_start=1  
/product="ACO protease"  
/protein\_id="AAG3354.1"  
/db\_xref="GI:11244759"  
/translation="MWLLITLSFLASTAAQDGLKLEBDECAHPQWQVALYERGR  
FNGAGSLIPHWLVSAAHCRSMRVRLGELHNRKRDGPQLRTTRVIPPHYEARS  
HRNDIMLLRLVQPARLNPQVRPAVLPTRCFHPGECACVSGWGLVPLSSPVSLPDTLHC





Qy	61	TCCCTCGCTAAATG	CAGGAGAGGAGG	CAGAAACACAGGGAA	-GAGGAAGGGTGGG	1119
Db	142488	TCCCTCGCTAAATG	CAGGAGAGGAGG	CAGAAACACAGGGAA	AGGAGGAGGAGG	142429
Qy	120	GAAGAAAGAGAGAG	GAGAGAGAGAGAG	CACAGATTAACACAACT	TACAGAAACACAGAGAGACA	179
Db	142428	GAAGAAAGAGAGAG	GAGAGAGAGAGAG	CACAGATTAACACAACT	TACAGAAACACAGAGAGACA	142369
Qy	180	CACAGAGAGCTGGG	CACAGAGGACACAC	CAGAGTCAGAGAGAAA	GAAGAGATAGAGAAA	239
Db	142368	CACAGAGAGCTGGG	CACAGGGACACAC	CAGAGTCAGAGAGAAA	GAAGAGATAGAGAAA	142309
Qy	240	GACACAAATGGAG	CACACAGAGGTG	TAAAGAAAGAGAGAT	TAAACAGAGTCCCGATACACG	299
Db	142308	GACACAAATGGAG	CACACAGAGGTG	TAAAGAAAGAGAGAT	TAAACAGAGTCCCGATACACG	142249
Qy	300	CAAGGGGCGAAGCA	CAGATTTTTCAGGGT	GTGCTATGATCATCT	TTCTTTTTTTTTTT	359
Db	142248	CAAGGGGCGAAGCA	CAGATTTTTCAGGGT	GTGCTATGATCATCT	TTCTTTTTTTTTTT	142189
Qy	360	TTTTTTTTTTTTTTT	TCAGACGGAGTCT	CGCTCTGTCGCCAGGCT	GCAGTGCAGTGGCG	419
Db	142188	TTTTTTTTTTTTTTT	TCAGACGGAGTCT	CGCTCTGTCGCCAGGCT	GCAGTGCAGTGGCG	142129
Qy	420	GGATCTCGGCTCA	CTGCAAGCTCCGCT	TCCGGGTTCACGCCAT	TTCTCCTCGCTCAGCCT	479
Db	142128	GGATCTCGGCTCA	CTGCAAGCTCCGCT	TCCGGGTTCACGCCAT	TTCTCCTCGCTCAGCCT	142069
Qy	480	CCCAAGTAGCTGG	AGCTACAGGCGCCG	CCCACTACGCGCCGGCT	TAATTTTTTTT	539
Db	142068	CCCAAGTAGCTGG	AGCTACAGGCGCCG	CCCACTACGCGCCGGCT	TAATTTTTTTT	142009
Qy	540	AGTAGACAGGGT	TTTTCACCGTTTT	TAGCGGGATGCTCGAT	CTCTGACCTCGTGATC	599
Db	142008	AGTAGACAGGGT	TTTTCACCGTTTT	TAGCGGGATGCTCGAT	CTCTGACCTCGTGATC	141949
Qy	600	CGCCCGCTCGGCT	CCCAAAGTCTGGG	ATTACAGCGTGAGCAC	CCGCGCCCGGCCAT	659
Db	141948	CGCCCGCTCGGCT	CCCAAAGTCTGGG	ATTACAGCGTGAGCAC	CCGCGCCCGGCCAT	141889
Qy	660	GATCATCTCTTG	ACTATGCTGATG	TGACAGATTAACAG	CCATACAGCTTACCTTT	719
Db	141888	GATCATCTCTTG	ACTATGCTGATG	TGACAGATTAACAG	CCATACAGCTTACCTTT	141829
Qy	720	AAATATCAGTTT	GGGCGCAGCCG	CTGCTGTAATTCAG	ACCTTTGGGAGG	779
Db	141828	AAATATCAGTTT	GGGCGCAGCCG	CTGCTGTAATTCAG	ACCTTTGGGAGG	141769
Qy	780	CAGAGGTGGGTGA	ATCATCTTGAGCC	AGAGTTTGAGACCAG	CGCTTGGCCAACTGTTGA	839
Db	141768	CAGAGGTGGGTGA	ATCATCTTGAGCC	AGAGTTTGAGACCAG	CGCTTGGCCAACTGTTGA	141709
Qy	840	ACTCTGCTTTT	TACTAAAAA	AAAAAATAAATAA	AAATCAGCGGGTGTCTGGGCA	899
Db	141708	ACTCTGCTTTT	TACTAAAAA	AAAAAATAAATAA	AAATCAGCGGGTGTCTGGGCA	141649
Qy	900	CACCTGTAATCC	CAAGCTATGCTG	GAGGCTGAGGCA	CGAGAGTCACTTGAACCTTGGAGG	959
Db	141648	CACCTGTAATCC	CAAGCTATGCTG	GAGGCTGAGGCA	CGAGAGTCACTTGAACCTTGGAGG	141589
Qy	960	GGAGGTTGCAGT	GGGCGGAGATCA	CATCACC	CGCCCTCCAGCCTTGGCGCAGAGCAAGAC	1019
Db	141588	GGAGGTTGCAGT	GGGCGGAGATCA	CATCACC	CGCCCTCCAGCCTTGGCGCAGAGCAAGAC	141529
Qy	1020	TCGTCTCAATA	TAATAAATAA	CAAAAGAGAGG	TTTGTGACCTTAGTTATATC	1079
Db	141528	TCGTCTCAATA	TAATAAATAA	CAAAAGAGAGG	TTTGTGACCTTAGTTATATC	141469
Qy	1080	TAAAAAATAAT	CTGTCACAAAT	TAGCAGAGAGT	GAAATAAGGAAATAAATGGCC	1139
Db	141468	TAAAAAATAAT	CTGTCACAAAT	TAGCAGAGAGT	GAAATAAGGAAATAAATGGCC	141409
Qy	1140	AAGAACTTAAG	GTATATTTTGA	CAAAATCATTTAG	AAATCTTTAAAAAAGAAATCAAG	1199
Db	141408	AAGAACTTAAG	GTATATTTTGA	CAAAATCATTTAG	AAATCTTTAAAAAAGAAATCAAG	141349
Qy	1200	AGGCATAGAA	AGACAGGAGAC	AGGAGACAGAAAC	CTCTGGCCCAAGGAGACAA	1259
Db	141348	AGGCATAGAA	AGACAGGAGAC	AGGAGACAGAAAC	CTCTGGCCCAAGGAGACAA	141289
Qy	1260	AACAAGGCTCCT	TAAGACAGAC	AGGAGAGAGAGAG	AGTGTAGTGAGAGACAGACAGA	1319
Db	141288	AACAAGGCTCCT	TAAGACAGAC	AGGAGAGAGAGAG	AGTGTAGTGAGAGACAGACAGA	141229
Qy	1320	GAANAAGACAG	AGAGAGAGAGAC	CAGAGACAGAGAG	CGCGAGAGGATAGAAAG	1379
Db	141228	GAANAAGACAG	AGAGAGAGAGAC	CAGAGACAGAGAG	CGCGAGAGGATAGAAAG	141169
Qy	1380	AGAGAGAGGGT	GCAGAGAGAC	CAGAGAGATTTG	GAGAGACTCAGAAAGATAGCCGAGGG	1439
Db	141168	AGAGAGAGGGT	GCAGAGAGAC	CAGAGAGATTTG	GAGAGACTCAGAAAGATAGCCGAGGG	141109
Qy	1440	AGAACACACAG	AGAGATGGAAGAG	ACTCTGAGAAAAA	CCAGAGACAAACATGGAAGAG	1499
Db	141108	AGAACACACAG	AGAGATGGAAGAG	ACTCTGAGAAAAA	CCAGAGACAAACATGGAAGAG	141049
Qy	1500	GAGTATCGAGG	GTAAACAGACAG	TGTTGGAATGAG	CAAAATGAGAGAGAAACAGCA	1559
Db	141048	GAGTATCGAGG	GTAAACAGACAG	TGTTGGAATGAG	CAAAATGAGAGAGAAACAGCA	140989
Qy	1560	ATCCAGGCGCC	CAAGATAGTACC	CCAGAGTTTGT	GAGAACCCAGATCTTTAAGGCTGGG	1619
Db	140988	ATCCAGGCGCC	CAAGATAGTACC	CCAGAGTTTGT	GAGAACCCAGATCTTTAAGGCTGGG	140929
Qy	1620	GAGCAGGGAAG	GGGCTGCGCT	TTCGGAGACCC	CTCCCATTTCTCCGGCGCAGGGA	1679
Db	140928	GAGCAGGGAAG	GGGCTGCGCT	TTCGGAGACCC	CTCCCATTTCTCCGGCGCAGGGA	140869
Qy	1680	GGTAGGAGTGA	CATTCGGAGCT	TGGGTGGGGT	GTCTGGGGGTGGAGATAGGGGAGC	1739
Db	140868	GGTAGGAGTGA	CATTCGGAGCT	TGGGTGGGGT	GTCTGGGGGTGGAGATAGGGGAGC	140809
Qy	1740	AGGAGAGCTAT	TGCTAAGGCCCG	ATAGGCACTCAT	TGCCCCGGGAATGTGCCCCAGGGA	1799
Db	140808	AGGAGAGCTAT	TGCTAAGGCCCG	ATAGGCACTCAT	TGCCCCGGGAATGTGCCCCAGGGA	140749
Qy	1800	GCACTGGCTGG	TTTATAAATC	CAGGCCCGGT	GCAGAGCCCGAGGAGGAGTGGCCAGG	1859
Db	140748	GCACTGGCTGG	TTTATAAATC	CAGGCCCGGT	GCAGAGCCCGAGGAGGAGTGGCCAGG	140689
Qy	1860	AAGGCAAGGCT	CTGAGAGTCT	CGGCTGAGCT	GGGAGCAAAATCCCCCACTACCTG	1919
Db	140628	AAGGCAAGGCT	CTGAGAGTCT	CGGCTGAGCT	GGGAGCAAAATCCCCCACTACCTG	140629
Qy	1920	GGGACAGGGCA	AGTGAAGCTT	GGTGGGTGGCT	CAGCAGGCGGAGAGAGGTGTC	1979
Db	140628	GGGACAGGGCA	AGTGAAGCTT	GGTGGGTGGCT	CAGCAGGCGGAGAGAGGTGTC	140569
Qy	1980	TGTGCGCTCTG	CAACCCACAT	CTTCTGTCCCT	CTTGTCCCTGTCTGGAGGCTGCTAG	2039
Db	140568	TGTGCGCTCTG	CAACCCACAT	CTTCTGTCCCT	CTTGTCCCTGTCTGGAGGCTGCTAG	140509
Qy	2040	ACTCTATCTT	CTGAATTTCT	ATAGTCCCT	TGGGTTCTCAGCGCAGTCCCGATGGTGGCCCGT	2099
Db	140508	ACTCTATCTT	CTGAATTTCT	ATAGTCCCT	TGGGTTCTCAGCGCAGTCCCGATGGTGGCCCGT	140449
Qy	2100	CCTTGTGGTTC	CTCTACTCTGGG	GAATAAGT	TAGGGAGGAGGGGAAGTGGTTAAG	2159
Db	140448	CCTTGTGGTTC	CTCTACTCTGGG	GAATAAGT	TAGGGAGGAGGGGAAGTGGTTAAG	140389
Qy	2160	GGCTCCCCGG	ATCGCTGGCC	TCCCAA	CCCTCTGACATTTCCCATTCACAGTGCAGCGGC	2219
Db	140388	GGCTCCCCGG	ATCGCTGGCC	TCCCAA	CCCTCTGACATTTCCCATTCACAGTGCAGCGGC	140329
Qy	2220	CATGGCTAC	AGCAAGAC	CCCCCTCTGG	ATGTGGGTGCTCTGTGCTCTGATCAGCCTTGCT	2279



Db	140328	CATGGCTACAGCAGACCCCTCGATGTGGTGCTCTGTGCTCTGATCAAGCCCTTGCT	140269	Db	139248	AGAACGAGCGCGCGGTGGCTCAGCCCTGTAATCCAGCCCTTTGGAGGCGCGAGGC	139189
Qy	2280	TTGCGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGATTGGGAGGACTG	2339	Qy	3360	AGCGGATCACTAGGTCCGGAGATCGAGGCCAGCCTGACCAACATGAGGAAACCCCGTC	3419
Db	140268	TTGCGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGATTGGGAGGACTG	140209	Db	139188	AGCGGATCACTAGGTCCGGAGATCGAGGCCAGCCTGACCAACATGAGGAAACCCCGTC	139129
Qy	2340	TTCTCGGCGCACTAGAGCGCTGTCCCTCGGGAACCTGTGAGCCTGGGCATGACTCCG	2399	Qy	3420	TTCTACTAAATAACAAAAATTAGCTGTGTGGCGCTGCTGTAAATCCCAAGCTACT	3479
Db	140208	TTCTCGGCGCACTAGAGCGCTGTCCCTCGGGAACCTGTGAGCCTGGGCATGACTCCG	140149	Db	139128	TTCTACTAAATAACAAAAATTAGCTGTGTGGCGCTGCTGTAAATCCCAAGCTACT	139069
Qy	2400	GGACCGGGTGAATGTGAGTCTCTGTCTGTACTTGTGTGTGGATGCTATGTGGCCCTG	2459	Qy	3480	CGGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGGAGGTGGAGTTGCCGTGAGCCGAG	3539
Db	140148	GGACCGGGTGAATGTGAGTCTCTGTCTGTACTTGTGTGTGGATGCTATGTGGCCCTG	140089	Db	139068	CGGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGGAGGTGGAGTTGCCGTGAGCCGAG	139009
Qy	2460	TGACTGCCACGGTGTGTGTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGGC	2519	Qy	3540	ATCGCGCATTTGCACTCCAGCTGGGGCAACAGAGCGAAACTCTCGTCTCGAAGAAAAAA	3599
Db	140088	TGACTGCCACGGTGTGTGTGGGAGGGGATGCTTTTCCCATATCAGGTGACTGTGGC	140029	Db	139008	ATCGCGCATTTGCACTCCAGCTGGGGCAACAGAGCGAAACTCTCGTCTCGAAGAAAAAA	138949
Qy	2520	GCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGTGATTTAA	2579	Qy	3600	AGAAAAAAAAGGGTAAGAACCAAGTGAATGGGCACGGGAGGACTGATGATGGAGTGGG	3659
Db	140028	GCAGGTGGCACTGACCCCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGTGATTTAA	139969	Db	138948	AGAAAAAAAAGGGTAAGAACCAAGTGAATGGGCACGGGAGGACTGATGATGGAGTGGG	138889
Qy	2580	GATTGTGTGGTCCCAAGCTGTGGGTGAATGCAATGAGCACTGGGGTGTTCACCTG	2639	Qy	3660	CATGCATGTAGTCTGTAGGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAAGCA	3719
Db	139968	GATTGTGTGGTCCCAAGCTGTGGGTGAATGCAATGAGCACTGGGGTGTTCACCTG	139909	Db	138888	CATGCATGTAGTCTGTAGGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAAGCA	138829
Qy	2640	TGTGTTTGGTGTGTGGTGAATTTGATATGATGACTGCAAGGTATCTGCAAGTTTCT	2699	Qy	3720	TGTTTTTCATCTGAGAAATTCAGAAACCTAGCCCTGTCTTCCCTCCATGTGGCCCCCTAA	3779
Db	139908	TGTGTTTGGTGTGTGGTGAATTTGATATGATGACTGCAAGGTATCTGCAAGTTTCT	139849	Db	138828	TGTTTTTCATCTGAGAAATTCAGAAACCTAGCCCTGTCTTCCCTCCATGTGGCCCCCTAA	138769
Qy	2700	GTCCCTGAGTCCCGGATTTGCGTGAACAAAGTGTCTATCACCATTGGAAGCTGTGAC	2759	Qy	3780	GCTGAGCCCTTCTTCTGCTGCTGCTTTCGGAAACCTAGCTCGGCCCATGAGCTCTGAC	3839
Db	139848	GTCCCTGAGTCCCGGATTTGCGTGAACAAAGTGTCTATCACCATTGGAAGCTGTGAC	139789	Db	138768	GCTGAGCCCTTCTTCTGCTGCTGCTTTCGGAAACCTAGCTCGGCCCATGAGCTCTGAC	138709
Qy	2760	TGTGTGCTGTGTCAGCGCAATTTGATTTGCTGAGTGTGAGCTTATGATGACCCCT	2819	Qy	3840	CCACACTCTCTTCTCAACACACGCCCTAGGCCAGACTCTAGTGGACCCGCCCTTAAGGCC	3899
Db	139788	TGTGTGCTGTGTCAGCGCAATTTGATTTGCTGAGTGTGAGCTTATGATGACCCCT	139729	Db	138708	CCACACTCTCTTCTCAACACACGCCCTAGGCCAGACTCTAGTGGACCCGCCCTTAAGGCC	138649
Qy	2820	ATTGTGACCGGTGACTACCTGAAGCTCTGTGAGGGGTGACTGTATGTGACTGTGTGT	2879	Qy	3900	ACACCCCTTTGGCCAGGCTCCACCCCTATTCTGTGGGTACCTTTAGAACCCCTTCA	3959
Db	139728	ATTGTGACCGGTGACTACCTGAAGCTCTGTGAGGGGTGACTGTATGTGACTGTGTGT	139669	Db	138648	ACACCCCTTTGGCCAGGCTCCACCCCTATTCTGTGGGTACCTTTAGAACCCCTTCA	138589
Qy	2880	GTCTGTGAGCGCGTGTAAATGCTACTGTATGTGTGATGAGTGTGAGCTGTGTGAG	2939	Qy	3960	AGTTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	4019
Db	139668	GTCTGTGAGCGCGTGTAAATGCTACTGTATGTGTGATGAGTGTGAGCTGTGTGAG	139609	Db	138588	AGTTCAGAGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	138529
Qy	2940	TTTCTGTCTCTGCTGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGGAGATGGTG	2999	Qy	4020	GTGAGTGGCGTCACTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGTATCTCG	4079
Db	139608	TTTCTGTCTCTGCTGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGGAGATGGTG	139549	Db	138528	GTGAGTGGCGTCACTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGTATCTCG	138469
Qy	3000	CCAGGTGACTGACTTGCAGTGTGTGCTGTGTGAGAGAGTATGTGGCAGTCTGAACAT	3059	Qy	4080	TGCTTCCACTCTCTGAGTGTGGGATTAACAGGTGGCGCCACACCGCTGGCTTAATTTT	4139
Db	139548	CCAGGTGACTGACTTGCAGTGTGTGCTGTGTGAGAGAGTATGTGGCAGTCTGAACAT	139489	Db	138468	TGCTTCCACTCTCTGAGTGTGGGATTAACAGGTGGCGCCACACCGCTGGCTTAATTTT	138409
Qy	3060	CTGTGCAACACGCGATCTGTGCTGTGCACTGAGACACTGTGGATGAGGGTGTGCGATCC	3119	Qy	4140	TGTGTCTTTAGTAGAGACAGGGTTTACCTTTTGGCCAGGCTGGTCTCAAACTCCCAAC	4199
Db	139488	CTGTGCAACACGCGATCTGTGCTGTGCACTGAGACACTGTGGATGAGGGTGTGCGATCC	139429	Db	138408	TGTGTCTTTAGTAGAGACAGGGTTTACCTTTTGGCCAGGCTGGTCTCAAACTCCCAAC	138349
Qy	3120	CGCTAGGCTCCCGGAGCGTGTACTCTGGAGACAGCTGTATGTAGCTGACCTGT	3179	Qy	4200	CTCAGGTGATCCGCCACCTCGGCTCCAGAGTGTGGGGTTACAGCGGTGAGCCACCG	4259
Db	139428	CGCTAGGCTCCCGGAGCGTGTACTCTGGAGACAGCTGTATGTAGCTGACCTGT	139369	Db	138348	CTCAGGTGATCCGCCACCTCGGCTCCAGAGTGTGGGGTTACAGCGGTGAGCCACCG	138289
Qy	3180	GGAGGCAACATGGCGGTGTCTGAGAACTGGTGGTGGTGGTGTACTGTGTTGTG	3239	Qy	4260	CCCCAGCCCCAAAGTCAGAGCTCTTTATAGGAGACTCTAACTGTAACCCCTGACCTGGC	4319
Db	139368	GGAGGCAACATGGCGGTGTCTGAGAACTGGTGGTGGTGGTGTACTGTGTTGTG	139309	Db	138288	CCCCAGCCCCAAAGTCAGAGCTCTTTATAGGAGACTCTAACTGTAACCCCTGACCTGGC	138229
Qy	3240	CGCGTGTCTTGGGGTGAAGTCTGTAATGATGTGTGGCCAGGCCCATCAGCAAGGGTA	3299	Qy	4320	CCTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGC	4379
Db	139308	CGCGTGTCTTGGGGTGAAGTCTGTAATGATGTGTGGCCAGGCCCATCAGCAAGGGTA	139249	Db	138228	CCTAACTAAGTCAATTCAAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGC	138169
Qy	3300	AGAACGAGCGCGCGGTGGCTCAGCCCTGTAATCCAGCCCTTTGGAGGCGCGAGGC	3359	Qy	4380	CTGACCCCACTTCTGAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCA	4439
				Db	138168	CTGACCCCACTTCTTGGAGCAAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCA	138109

QY	4440	GGCTCCAGCCCCACAGCTTTGGCACTTACCCCTGAGCTTGTCAGGAATCCTGTACCCAA	4499
Db	138108	GGCTCCAGCCCCACAGCTTTGGCACTTACCCCTGAGCTTGTCAGGAATCCTGTACCCAA	138049
QY	4500	TTTATCCCTCACATGATAGTCTTAGCCATTCACAGGAATCTGTGAGGTCCAGTTAGAGTCC	4559
Db	138048	TTTATCCCTCACATGATAGTCTTAGCCATTCACAGGAATCTGTGAGGTCCAGTTAGAGTCC	137989
QY	4560	AGTAACCTTACCTGAGCGCTGGCTCTGTCTCTGAGCTTGAGCTTGAGCGCTTGAGAGGTGCC	4619
Db	137988	AGTAACCTTACCTGAGCGCTGGCTCTGTCTCTGAGCTTGAGCTTGAGCGCTTGAGAGGTGCC	137929
QY	4620	ACTCTTATTCTCCAGGCCCTGCCCCCTGCCCCCTCAGCATGTCAGACACCCACCCTCTAGC	4679
Db	137928	ACTCTTATTCTCCAGGCCCTGCCCCCTGCCCCCTCAGCATGTCAGACACCCACCCTCTAGC	137869
QY	4680	TGCTCTGGCCCTCTGTAGTCTGAACCCACCCAGCCCAAGCCCGCTCTGAGCCCGCC	4739
Db	137868	TGCTCTGGCCCTCTGTAGTCTGAACCCACCCAGCCCAAGCCCGCTCTGAGCCCGCC	137809
QY	4740	CCAAACCATTTTCCGTTTCCAGAGCATGTTCTCGCAACAATCATGTTTCTCTGTGACCCAC	4799
Db	137808	CCAAACCATTTTCCGTTTCCAGAGCATGTTCTCGCAACAATCATGTTTCTCTGTGACCCAC	137749
QY	4800	CCCTCTAACACCGCTGCCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGCGCGGGAAGAC	4859
Db	137748	CCCTCTAACACCGCTGCCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGCGCGGGAAGAC	137689
QY	4860	GCCCGGTGCGATGACAGCAGCAGCCGCGCATCATCAATGATCCGACTGCGATATGCACACC	4919
Db	137688	GCCCGGTGCGATGACAGCAGCAGCCGCGCATCATCAATGATCCGACTGCGATATGCACACC	137629
QY	4920	CAGCGCTGGCAGCGCGCTGTTGCTAAGCCCAACAGCTTACTGCGGGCGGTGTTG	4979
Db	137628	CAGCGCTGGCAGCGCGCTGTTGCTAAGCCCAACAGCTTACTGCGGGCGGTGTTG	137569
QY	4980	GTGCTATCCACAGTGGGTGCTC 5000	
Db	137568	GTGCTATCCACAGTGGGTGCTC 137548	
RESULT 7			
AC027602/c			
LOCUS		217346 bp DNA linear HTG 24-AUG-2002	
DEFINITION		Homo sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT	
SEQUENCE		SEQUENCE, 12 unordered pieces.	
AC027602			
VERSION		AC027602.4 GI:11178143	
KEYWORDS		HTG; HTGS PHASE1; HTGS_DRAFT.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE		Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,	
AUTHORS		Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,	
TITLE		Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,	
JOURNAL		Campioano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,	
AUTHORS		Collins,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,	
		Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,	
		Grand-Pierre,N., Grant,G., Hago,B., Heaford,A., Horton,L.,	
		Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,	
		Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J.,	
		Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,	
		McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,	
		Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,	
		Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,	

O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 217346)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bada,F., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Burkett,G., Campioano,A., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hago,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 15, 2000 this sequence version replaced gi:11136831.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9166  
Center clone name: 795\_B.6  
----- Summary Statistics  
Sequencing vector: M13; M77815; 31% of reads  
Sequencing vector: Plasmid; n/a; 69% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 210748 bases at least Q40  
Consensus quality: 213655 bases at least Q30  
Insert size: 194000; agarose-fp  
Insert size: 216246; sum-of-contigs  
Quality coverage: 11.9 in Q20 bases; agarose-fp  
Quality coverage: 10.7 in Q20.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 8149: contig of 8149 bp in length  
\* 8150 8249: gap of 100 bp  
\* 8250 9592: contig of 1343 bp in length  
\* 9593 9692: gap of 100 bp  
\* 9693 10733: contig of 1041 bp in length  
\* 10734 10833: gap of 100 bp  
\* 10834 13519: contig of 2686 bp in length

*	13520	13619:	gap of 100 bp
*	13620	17510:	contig of 3891 bp in length
*	17511	17610:	gap of 100 bp
*	17611	24602:	contig of 6392 bp in length
*	24603	24702:	gap of 100 bp
*	24703	35434:	contig of 10732 bp in length
*	35435	35534:	gap of 100 bp
*	35535	124474:	contig of 88940 bp in length
*	124475	124574:	gap of 100 bp
*	124575	134664:	contig of 10090 bp in length
*	134665	134764:	gap of 100 bp
*	134765	162343:	contig of 27579 bp in length
*	162344	162443:	gap of 100 bp
*	162444	208917:	contig of 46474 bp in length
*	208918	209017:	gap of 100 bp
*	209018	217346:	contig of 8329 bp in length

## FEATURES

```

source
1. 217346
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="19"
  /map="19"
  /clone="RP11-795B6"
  /clone.lib="RP11 Human Male BAC"
misc feature
1. 8149

```

misc feature

```

/note="assembly_fragment
clone_end:SP6
vector_side:left"

```

misc feature	8250.79592
vector_size	

```
misc_feature 8250. .9392 /note="assembly_fragment"  
misc_feature 9693. .10733
```

```
misc_feature      /note="assembly_fragment"  
10834.  .13519
```

```
misc_feature      /note="assembly_fragment"  
13620. 17510
```

```
misc_feature
17611. .24602
/note="assembly_fragment"
/note="assembly_fragment"
```

```
misc_feature      24703.35434 /note="assembly_fragment"
```

```
misc_feature'
/note="assembly_fragment
35535.124474
/note="assembly_fragment"
```

```
misc_feature
124575. .134664
/note="assembly fragment"
```

```
misc_feature
134765. .162343
/note="assembly_fragment"
```

```
misc_feature 162444. .208917-
/note="assembly_fragment"
```

```
misc_feature      209018. .217346  
                  /note="assembly_fragment
```

```
clone_end:T7
```

## ORIGIN

Query Match	84.9%;	Score	4243.4;	DB 2;	Length	217346;
Best Local Similarity	99.0%;	Pred. NO.	0;			
Matches 4335;	Conservative	0;	Mismatches	11;	Indels	35;
					Gaps	5;

Qy	654	GGCCATGATCATCTTCTTGACTATGCTGATGTGA	CAAGTACCTAAAGCCATCAGACTCTA	713
Db	203789	GTCTATGATCATCTTCTTGACTATGCTGATGTGA	CAAGTACCTAAAGCCATCAGACTCTA	203730

Qy 714 CCCCTTTAAATATGCAAGTTTGGGCCAGGCA CCGTGGCTCATGCCTGTAAATTCACGACCTTT 773

Db 203729 CCCTTTAAATATGCAGTTTGGGCCAGGACCGTGGCTCATGCTGTAAATTCAGACTTT 203670

Qy 774 GGGAGGCAGAGGTGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGCCTGGCCAACT 833

Db 203669 GGGAGGCAGAGGTGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAAGCTGGCCCAACAT 203610

QY 834 GGTGAACCTCTGCTTTACTAAAAAATAAAAAAATACAGCCGGTGTCTGT 893

\_\_\_\_\_

D	b	203609	GGTGAAGACTCTGTCTTTACT-AAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCGT	203555
Q	y	894	GGGCACACCTGTAAATCCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCATCTTGAAACCGCT	953
D	b	203550	GGGCACACCTGTAAATCCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCATCTGNAACCGCT	203491
Q	y	954	GGAGCGGAGGTTGCAGTGGGCCGAGATCACATCACTCCGCCCTCCAGCCTCGGGCGACAGAG	1013
D	b	203490	GGAGCGGAGGTTGCAGTGGGCCGAGATCACATCACCGCCCTCCAGCCTCGGGCGACAGAG	203431
Q	y	1014	CAAGACTCTGTCTCAAATAAATAAACAACAACGAAACAGACAGTGTGTGCTTACCTTAGT	1073
D	b	203430	CAAGACTCTGTCTCAAATAAATAAACAACAACGAAACAGACAGTGTGTGCTTACCTTAGT	203371
Q	y	1074	TATATCTAAAAAAAATTCGTCTCAACAAATAGACAGAACTGAAATAAAGGAAAAATAAA	1133
D	b	203370	TATACCTTAAAAAAAATTCGTCTCAACAAATAGACAGAACTGAAATAAAGGAAAAATAAA	203311
Q	y	1134	TGGGCCAAGAATCTTAAGGTATATTGTGACAAATCATTCAGAACCTTTAAAAAAGAAAGAA	1193
D	b	203310	TGGGCCAAGAATCTTAAGGTATATTGTGACAAATCATTCAGAACCTTTAAAAAAGAAAGAA	203251
Q	y	1194	TCACAGAGGCATAGAAAGACAGGAGGACNACGGGAGACAGAAACACCTGTGGCCCCAAGA	1253
D	b	203250	TCACAGAGGCATAGAAAGACAGGAGGAGNACAGGAGACAGAAACACCTGTGGCCCCAAGA	203191
Q	y	1254	GAAACAAACAAGGCTCTTAAGACAGACAGAGAG---GAGACAGACAGAGAGTGTAGTGAGA	1309
D	b	203190	GAAACAAACAAGGCTCTTAAGACAGACAGAGAG---GAGACAGACAGAGAGTGTAGTGAGA	203131
Q	y	1310	GACAGACAGAGAAAAAGACAGAGAGAGAGACAGACAGACAGAGACAGAGAGGCGAGAG	1369
D	b	203130	GACAGACAGAGAAAAAGACAGAGAGAGAGACAGACAGACAGAGACAGAGAGGCGAGAG	203071
Q	y	1370	GGATAGAAAGACAGAGAGGGGTGCAGAGAGACACGAGATATTGACAGAGACTTCAGAAAGA	1429
D	b	203070	GGATAGAAAGACAGAGAGGGGTGCAGAGAGACACGAGATATTGACAGAGACTTCAGAAAGA	203011
Q	y	1430	TAGCCGAGGAGAACACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGACACAAG	1489
D	b	203010	TAGCCGAGGAGAACACACAGAGAGATGGAAGAAGACTCTGAGAAAAAACACAGACACAAG	202951
Q	y	1490	ATGAAAGAGGAGTATCCAGGGTGAAACAGACAGTGGTGGATGAGCAAAATGCGAGAGAG	1549
D	b	202950	ATGAAAGAGGAGTATCCAGGGTGAAACAGACAGTGGTGGATGAGCAAAATGCGAGAGAG	202891
Q	y	1550	AAAGCAAGCAATCCAGCGGCCAAGAAATAGTACCAGAGTTGGTGAGAGCCAGATCCTT	1609
D	b	202890	AAAGCAAGCAATCCAGCGGCCAAGAAATAGTACCAGAGTTGGTGAGAGCCAGATCCTT	202831
Q	y	1610	AAGSCTGGGGAGSCAGGAGGGCTTGGCTTCCGAGACCCCTCCCAATTCCTCC	1669
D	b	202830	AAGSCTGGGGAGSCAGGAGGGCTTGGCTTCCGAGACCCCTCCCAATTCCTCC	202771
Q	y	1670	GGGCCAGGAGGTAGGAGTGACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGGAGA	1729
D	b	202770	GGGCCAGGAGGTAGGAGTGACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGGAGA	202711
Q	y	1730	TAGGGGAGACAGAGAGACTATTGCTAAGGCCGATAGGCACCTCATTTGCCCGGGAATGT	1789
D	b	202710	TAGGGGAGACAGAGAGACTATTGCTAAGGCCGATAGGCACCTCATTTGCCCGGGAATGT	202651
Q	y	1790	GCCCGAGGAGCAGTGGGTGTTATAACTCAGGCCCGGTGCCAGAGCCACAGGAGGAGGC	1849
D	b	202650	GCCCGAGGAGCAGTGGGTGTTATAACTCAGGCCCGGTGCCAGAGCCACAGGAGGAGGC	202591
Q	y	1850	AGTGGCCAGGAAGGCACAGGCCTCAGAAAGTCTGCGGCTGAGCTGGGAGCAAAATCCCCAC	1909
D	b	202590	AGTGGCCAGGAAGGCACAGGCCTCAGAAAGTCTGCGGCTGAGCTGGGAGCAAAATCCCCAC	202531
Q	y	1910	CCCCTACTGGGGACAGGGCAAGTGAGACTGTGTAGGGGTGGCTCAGCAGGCGAGGAG	1969
D	b	202530	CCCCTACTGGGGACAGGGCAAGTGAGACTGTGTAGGGGTGGCTCAGCAGGCGAGGAG	202471

QY	1970	GAGAGGTGTGTGCGCTCTGQACCAACATCTTTCTCTGTGCCCTCCTTCCCTGTCTGG	2029	3021	TGTGCTGTGTGCAGAGAAGATATGTGGCAGTCTGAACATCTGTGCACACACGCGCATCTGT	3080
Db	202470	GAGAGGTGTGTGTGCGCTCTGQACCAACATCTTTCTCTGTGCCCTCCTTCCCTGTCTGG	202411	201390	TGTGCTGTGTGCAGAGAAGATATGTGGCAGTCTGAACATCTGTGCACACACGCGCATCTGT	201331
QY	2030	AGGCTGTAGACTCTCTCTCTGAAATCTATAGTGCTGGGTCTCAGCGCAGTGGCGAT	2089	3081	GGGTGCACCTGAGACACTGTGGATGAGGGTGTGCGATCCGGCTAGGCTGCCCGGGAGCGT	3140
Db	202410	AGGCTGTAGACTCTCTCTCTGAAATCTATAGTGCTGGGTCTCAGCGCAGTGGCGAT	202351	201330	GGGTGCACCTGAGACACTGTGGATGAGGGTGTGCGATCCGGCTAGGCTGCCCGGGAGCGT	201271
QY	2090	GGTGGCCCGTCTGTGTGGTTCCTCTACCTGGGGAATAAGGTAGGGGAGGGAGGGAA	2149	3141	GTGTACTCTGGAGACAGAGCTGTATGTAGTGCACCTGTGTGGAGGCAACATGGCGTGTCT	3200
Db	202350	GGTGGCCCGTCTGTGTGGTTCCTCTACCTGGGGAATAAGGTAGGGGAGGGAGGGAA	202291	201270	GTGTACTCTGGAGACAGAGCTGTATGTAGTGCACCTGTGTGGAGGCAACATGGCGTGTCT	201211
QY	2150	GTGGGTAAAGGCTCCCGGATCGCTGGCCCTCCCAACCTCTGACATTTCCCATCCAG	2209	3201	GCAGAACTCCGTGCTGTGGCTGTACTGTCTGTGTGGCGTGGTCTTTGGGGTGAAT	3260
Db	202290	GTGGGTAAAGGCTCCCGGATCGCTGGCCCTCCCAACCTCTGACATTTCCCATCCAG	202231	201210	GCAGAACTCCGTGCTGTGGCTGTACTGTCTGTGTGGCGTGGTCTTTGGGGTGAAT	201151
QY	2210	GTGAGGGCCATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTGATCA	2269	3261	TGCTGAATGATGGTGTGTGCAGGGCCATCAGACNAGGGTAAGAACAGGCCGGCGGTG	3320
Db	202230	GTGAGGGCCATGGCTACAGCAAGACCCCTCGATGTGGGTGCTCTGTGCTGATCA	202171	201150	TGCTGAATGATGGTGTGTGCAGGGCCATCAGCAAGGGTAAGAACAGGCCGGCGGTG	201091
QY	2270	CAGCTTGCTTCTGTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGGAT	2329	3321	GCTCAGCCTGTAAATCCAGCCCTTTGGGAGGCCGAGGAGCGGATCACCTCAGGTCGG	3380
Db	202170	CAGCTTGCTTCTGTGGGGTCAACAGGTAAACAGAACTCTGGGGTGGAGGGTGTGGGAT	202111	201090	GCTCAGCCTGTAAATCCAGCCCTTTGGGAGGCCGAGGAGCGGATCACCTCAGGTCGG	201031
QY	2330	GGGAGGACTGTCTCTGCGGCACTAGAGCGCTGTGCCCTGGGAACTGTGTGAGCCTGGG	2389	3381	GAGATCGAGGCCAGCCCTGACCAACATGGAGAACCCGCTCTACTATAAAATACAAAAAT	3440
Db	202110	GGGAGGACTGTCTCTGCGGCACTAGAGCGCTGTGCCCTGGGAACTGTGTGAGCCTGGG	202051	201030	GAGATCGAGGCCAGCCCTGACCAACATGGAGAACCCGCTCTACTATAAAATACAAAAAT	200971
QY	2390	CATGACTCCGGGACCGGCTGAATGTGAGTCTCTGTCTGTAATCCAGCTACTCGGGAGAGT	2449	3441	TAGCTGGTGTGGTGGCGTGTCTGTAAATCCAGCTACTCGGGAGAGTCTGGGGCAGAAAA	3500
Db	202050	CATGACTCCGGGACCGGCTGAATGTGAGTCTCTGTCTGTAATCCAGCTACTCGGGAGAGT	201991	200970	TAGCTGGTGTGGTGGCGTGTCTGTAAATCCAGCTACTCGGGAGAGTCTGGGGCAGAAAA	200911
QY	2450	TGTGGCCCTGTGACTGCCAGGTGTGTGCGGGAGGGGATGCCCTTTCCCATATCAGG	2509	3501	TGCTTTGAACCCGGGAGGTGGAGGTTGCGGTGAGCGCGAGATCGCGCATTTGCGACTCAGC	3560
Db	201990	TGTGGCCCTGTGACTGCCAGGTGTGTGCGGGAGGGGATGCCCTTTCCCATATCAGG	201931	200910	TGCTTTGAACCCGGGAGGTGGAGGTTGCGGTGAGCGCGAGATCGCGCATTTGCGACTCAGC	200851
QY	2510	TGACTGTGCGGCAAGTGGCACTGACCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGT	2569	3561	CTGGGCAACAGAGCGAAACTCCGCTCTCGAAAGAAAAAAGAAAAAGGTAAGAA	3620
Db	201930	TGACTGTGCGGCAAGTGGCACTGACCTTTGAGGCTGTGTGTGGTGTGTGATTTGTGT	201871	200850	CTGGGCAACAGAGCGAAACTCCGCTCTCGAAAGAAAAAAGAAAAAGGTAAGAA	200791
QY	2570	GTGCATTTAAGATGTGTGGCTTCCACAGCTGTGGGTGAATGATGTAGCACTGGGG	2629	3621	CAAGTGAATGGGCACGGGAGGACTGATGATGAGTGGGGCATGATGTAGTCTGTAGGTC	3680
Db	201870	GTGCATTTAAGATGTGTGGCTTCCACAGCTGTGGGTGAATGATGTAGCACTGGGG	201811	200790	CAAGTGAATGGGCACGGGAGGACTGATGATGAGTGGGGCATGATGTAGTCTGTAGGTC	200731
QY	2630	GTGTTCACTGTGTGTGGCTGTGTGGTGTGTGCACTGGCAATGTATGACTGACG	2684	3681	TGTGTGTAGAGAGGAGATTGACAGGATTGAGAGGCAATGTTTTTCTGTGAGAAATTCAG	3740
Db	201810	GTGTTCACTGTGTGTGGCTGTGTGGTGTGTGCACTGGCAATGTATGACTGACG	201751	200730	TGTGTGTAGAGAGGAGATTGACAGGATTGAGAGGCAATGTTTTTCTGTGAGAAATTCAG	200671
QY	2685	-----GTATCTGCG-----TTCTCTGCTCCCTGAGGTCCTGGGATG	2720	3741	AAACCTAGGCTGTCTCTTCCCTCCATGTGGGCCCTTAAGCTGAGGCCCTCTTTTCTGTGT	3800
Db	201750	ATACATATCCCTGTATATATGAGGATACATATACATATCCCTGAGGTCCTGGGATG	201691	200670	AAACCTAGGCTGTCTCTTCCCTCCATGTGGGCCCTTAAGCTGAGGCCCTCTTTTCTGTGT	200611
QY	2721	CGTGCAACAAAGTGTGATCAACATGGAAGCTGTGATGTGTGTGTGTGTGTGTGTGT	2780	3801	CTGTCTTTTGGAACTCTAGCTCGGCCCATGAGCTGTGAACCCACCTCTTTTCTCAACCA	3860
Db	201690	CGTGCAACAAAGTGTGATCAACATGGAAGCTGTGATGTGTGTGTGTGTGTGTGTGT	201631	200610	CTGTCTTTTGGAACTCTAGCTCGGCCCATGAGCTGTGAACCCACCTCTTTTCTCAACCA	200551
QY	2781	TATGTGATTTGGCTGAGTGTGACGTTATGGATCCCGTATTTGTGACCGGTGACTACC	2840	3861	CGCCCCCTAGGCGAGACTCTAGTGGAGCCCCGCTTAAGGCCACACCCCTTTGGGGCAGGCTC	3920
Db	201630	TATGTGATTTGGCTGAGTGTGACGTTATGGATCCCGTATTTGTGACCGGTGACTACC	201571	200550	CGCCCCCTAGGCGAGACTCTAGTGGAGCCCCGCTTAAGGCCACACCCCTTTGGGGCAGGCTC	200491
QY	2841	TGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTGTGTAA	2900	3921	CACCCCTATTCTGTGGGTACCTTCTAGAAACCCCTTTCAAAGTTCAGAGC-TTTTTTTTTTT	3979
Db	201570	TGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTGTGTAA	201511	200490	CACCCCTATTCTGTGGGTACCTTCTAGAAACCCCTTTCAAAGTTCAGAGCTTTTTTTTTTT	200431
QY	2901	TGCTATGT	2960	3980	TTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGGTGTGAGTGTGAGTGTGATCTCGG	4039
Db	201510	TGCTATGT	201451	200430	TTTTTTTTTGGAGACAGTCTTGTCTCTCTCCAGGGTGTGAGTGTGAGTGTGATCTCGG	200371
QY	2961	ATAGAGGTGACGGGT	3020	4040	CTCAGTGCACCTCTGCGCTCCAGGTTCAAGTGAATTTCTGTGCTCCACCTCTCTGTAGTAG	4099
Db	201450	ATAGAGGTGACGGGT	201391	200370	CTCAGTGCACCTCTGCGCTCCAGGTTCAAGTGAATTTCTGTGCTCCACCTCTCTGTAGTAG	200311
				4100	CTGGGATTAAGGTGTGGCGCCACAGCGCTGGCTAATTTTTTGTGTCTTTAGTAGAGACAG	4159

Db 200310 CTGGATTACAGGTGCGCGCCACACCGCTGGCTTAATTTTGTGCTTTAGTAGACAG 200251  
QY 4160 GGTTTACCTTTGGCCAGCGTGTCTAAACTCCCAACTCAGGTGATCCGCCACCT 4219  
Db 200250 GGTTTACCTTTGGCCAGCGTGTCTAAACTCCCAACTCAGGTGATCCGCCACCT 200191  
QY 4220 CGGCTTCCCAAGTGTCTGGGGTTACAGCGTGTAGCGCAACCGCCGCCAGCCCAAGTACAG 4279  
Db 200190 CGGCTTCCCAAGTGTCTGGGGTTACAGCGTGTAGCGCAACCGCCGCCAGCCCAAGTACAG 200131  
QY 4280 CTCTTTATAGAGACTCTAATAGTAACTGACCTGACCTGGCCCTTAATAGTCAATCCAA 4339  
Db 200130 CTCTTTATAGAGACTCTAATAGTAACTGACCTGACCTGGCCCTTAATAGTCAATCCAA 200071  
QY 4340 ACCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCTGACCCCACTTCTTGAGAC 4399  
Db 200070 ACCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCTGACCCCACTTCTTGAGAC 200011  
QY 4400 CAGTTCATCCCTAAGCCCTGCTCTCCCTCCATCCCGAGGCTCCAGCCGCCACAGCTT 4459  
Db 200010 CAGTTCATCCCTAAGCCCTGCTCTCCCTCCATCCCGAGGCTCCAGCCGCCACAGCTT 199951  
QY 4460 TGGCACTACCCCTGAGCTTCTCAGGAATCTGTACCAATTTTACCTCACAATGAGTT 4519  
Db 199950 TGGCACTACCCCTGAGCTTCTCAGGAATCTGTACCAATTTTACCTCACAATGAGTT 199891  
QY 4520 CTAGCCCAATTCAGGAATCTGTGAGGTCCAGTGTAGAGTCCAGTAACCTTACCTGAGCCTG 4579  
Db 199890 CTAGCCCAATTCAGGAATCTGTGAGGTCCAGTGTAGAGTCCAGTAACCTTACCTGAGCCTG 199831  
QY 4580 GGCTCTGCTCTGAGCTTGTAGCTGGCTTGTAGAGGTGACACTTATTTCTCAGGCGCT 4639  
Db 199830 GGCTCTGCTCTGAGCTTGTAGCTGGCTTGTAGAGGTGACACTTATTTCTCAGGCGCT 199771  
QY 4640 GCCCTGCCCCCTCAGCATCTCAGACACCCACCTCTAGCTGGTCTGGCTCTTTGAGTCT 4699  
Db 199770 GCCCTGCCCCCTCAGCATCTCAGACACCCACCTCTAGCTGGTCTGGCTCTTTGAGTCT 199711  
QY 4700 GAAACCCACCCCGAGCCCAAGCCCGCTCTGAGCCCGCCCAACCCATTTTCGTTCCC 4759  
Db 199710 GAAACCCACCCCGAGCCCAAGCCCGCTCTGAGCCCGCCCAACCCATTTTCGTTCCC 199651  
QY 4760 AGAGCATGTTCTGCCAACAATGATGTTCTGTGACCACTCTTAACACCGTGCCCTC 4819  
Db 199650 AGAGCATGTTCTGCCAACAATGATGTTCTGTGACCACTCTTAACACCGTGCCCTC 199591  
QY 4820 TGGAGCAACACGAGCTGGAGCTGGGCGCGGAGAGCGCGGTGCGATGACAGCAG 4879  
Db 199590 TGGAGCAACACGAGCTGGAGCTGGGCGCGGAGAGCGCGGTGCGATGACAGCAG 199531  
QY 4880 CAGCCGATCATCAATGATCCGACTCGCATATGCACACCCAGCCGCGGAGCGCGCT 4939  
Db 199530 CAGCCGATCATCAATGATCCGACTCGCATATGCACACCCAGCCGCGGAGCGCGCT 199471  
QY 4940 GTTGCTAAGGCCCAACAGCTTCTACTCGGGGGGTTGTTGGTGATCCACAGTGGCTGCT 4999  
Db 199470 GTTGCTAAGGCCCAACAGCTTCTACTCGGGGGGTTGTTGGTGATCCACAGTGGCTGCT 199411  
QY 5000 C 5000  
Db 199410 C 199410

RESULT 8  
AC130782/c  
LOCUS Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12  
DEFINITION ordered pieces.  
ACCESSION AC130782  
VERSION AC130782.2 GI:25167101  
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.  
SOURCE Pan troglodytes (chimpanzee)

## ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

## REFERENCE

AUTHORS

1 (bases 1 to 200792)

Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Carliaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,  
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,  
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., McDowell, J.,  
Margulis, E.H., Masiello, C., Maskeri, B., Prasadi, A.,  
Paguirian, C., Pearson, R., Portnoy, M.E., Prasad, A.,  
Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,  
Stantibop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,  
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

NISC Comparative Sequencing Initiative

## TITLE

Unpublished

## JOURNAL

2 (bases 1 to 200792)

## REFERENCE

AUTHORS

Green, E.D.

## TITLE

Direct Submission

## JOURNAL

Submitted (14-AUG-2002)

## REFERENCE

AUTHORS

Submitted (14-AUG-2002)

## TITLE

Genome Center

## JOURNAL

Submitted (22-NOV-2002)

## REFERENCE

AUTHORS

Submitted (22-NOV-2002)

## TITLE

Genome Center

## JOURNAL

Submitted (22-NOV-2002)

## COMMENT

On Nov 22, 2002 this sequence version replaced gi:22218452.

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc\_zoo@nhgri.nih.gov

----- Project Information

Center project name: dhz

Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 196656 bases at least Q40  
Consensus quality: 197883 bases at least Q30  
Consensus quality: 198879 bases at least Q20  
Insert size: 215000; agarose-fp  
Insert size: 196692; sum-of-contigs  
Quality coverage: 9.03x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 28306: contig of 28306 bp in length  
\* 28307 28406: gap of unknown length  
\* 28407 37856: contig of 9450 bp in length  
\* 37857 37956: gap of unknown length

*	37957	73522:	cont:ig	of	35566	bp	in	length
*	73223	73622:	gap	of	unknown	length		
*	73283	83567:	cont:ig	of	9945	bp	in	length
*	83668	83667:	gap	of	unknown	length		
*	83568	88817:	cont:ig	of	5150	bp	in	length
*	88918	88917:	gap	of	unknown	length		
*	88918	85611:	cont:ig	of	36694	bp	in	length
*	125512	125711:	gap	of	unknown	length		
*	125712	125879:	cont:ig	of	34168	bp	in	length
*	159880	159979:	gap	of	unknown	length		
*	159980	174698:	cont:ig	of	14719	bp	in	length
*	174659	174798:	gap	of	unknown	length		
*	174799	186382:	cont:ig	of	11584	bp	in	length
*	186383	186482:	gap	of	unknown	length		
*	186483	193344:	cont:ig	of	6862	bp	in	length
*	193345	193444:	gap	of	unknown	length		
*	193445	199363:	cont:ig	of	5919	bp	in	length
*	199364	199463:	gap	of	unknown	length		
*	199464	200792:	cont:ig	of	1329	bp	in	length

Query Match	82.1%;	Score 4106.4;	DB 2;	Length 200792;
Best Local Similarity	97.3%;	Pred. No. 0;		
Matches 4268;	Conservative 0;	Mismatches 66;	Indels 52;	Gaps 7;
QY	654	GGCCATGATCATCTCTTTGACTATGCTGATGTCACAACTACCTAAAGCCATCAGACTCTA	713	
DB	78497	GTCATGATCATCTCTTTGACTATGCTGATGTCACAACTACCTAAAGCCATCAGACTCTA	78438	
QY	714	CCCTTTAAATATGCAGTTTGGCGAGCACCCGTGGCTCATGCTGTAAATCCAGCACTTT	773	
DB	78437	CCCTTTAAATATGCAGTTTGGCGAGCACCTGTGGCTCATGCTATAAATCCAGCACTTT	78378	
QY	774	GGGAGGCAGAGCTGGGTCAATCACTTCAGCCAGGAGTTTGAGACCCAGCGCTGCCAAACAT	833	
DB	78377	GGGAGGCAGAGCTGGGTGAATCACTTCAGCCAGGAGTTTGAGACCCAGCGCTGCCAAACAT	78318	
QY	834	GGTGAAACTCTGCTCTTTACTTAAAAAATAAAAAAAAAAAAAAAAAAATACAGCCGGGTGTGGT	893	

Db	78317	GGTGAAACTGTGCTTTACT-----AAAAAAAAAAAAAAAAAAAAATCAGCGGGTGTGCG	78264
Qy	894	GGGGCACACCTGTAATCCCAAGCTATGCTGGAGCGCTGAGGCACGAGAGTCACTTGAAACCT	953
Db	78263	GGGGCACACCTGTANTCCAGCTACGCCGAGAGGCTGAGGCACGAGAGTCATTGAAACCT	78204
Qy	954	GGAGGGGAGGTTGTCAGTGGGCGGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAG	1013
Db	78203	GGAGTGGAGTTGTCAGTGGGCGGAGATCACATCACCGCCCTCCAGCCTGGGCGACAGAG	78144
Qy	1014	CAGACTCTGCTCAAAATAAATAAATAAACAAGCAGAACTGTTGTACCTTGTAGT	1073
Db	78143	CAAGACTCTGCTCAAAATAAATAAATAAACAAGCAGAACTGTTGTACCTTGTAGT	78084
Qy	1074	TATATCT--AAAAAAAAAATGCTCAACAATAAGACAGAGTGAATAAAGGAAAAATA	1131
Db	78083	TATACCTAAAAAATAAATAAATAAATAAATAAAGAGAGTGAATAAAGGAAAAATA	78024
Qy	1132	AATGGGCCAAGAACTCTAAGGTATATTGACAAATCATTCAGAACTTTAAAAAAGAAAG	1191
Db	78023	AATGGGCCAAGAACTCTAAGGTATATTGACAAATCATTCAGAACTTTAAAAAAGAAAG	77964
Qy	1192	AATCACAGAGGCATAGAAAAGACAGGGAGGAAACAGGGAGACAGAAACACTGTGGCCCAAG	1251
Db	77963	AATCACAGAGGCATAGAAAAGACAGGGAGGAAACAGGGAGACAGAAACACTGTGGCCCAAG	77904
Qy	1252	GAGAACAAACACAGGCTCTTAGACAGACAGAG-- --GAGAGACAGAGAGAGTCAAGTGA	1307
Db	77903	GAGAACAAACACAGGCTCTTAGACAGACAGAGGCGAGAGAGAGAGAGTCAAGTGA	77844
Qy	1308	GAGACACACAGAAAAAGACAGACAGACAGACAGAG-----	1347
Db	77843	GAGACACACAGAAAAAGACAGACAGACAGAGAGACAGAGGCGAGGGATAGAAAG	77784
Qy	1348	-----CAGACAGACAGAGGCGAGGGATAGAAAAGACAGAGAGGGGTGA	1394
Db	77783	AGAGAGAGGGGTGGACAGAGACAGAGGCGAGAGGGATAGAAAAGACAGAGAGGGGTGA	77724
Qy	1395	GAGAGACACAGATATTGAGAGAGACTCAGAAGATAGCCGAGGAGAGAACACACAGAGAG	1454
Db	77723	GAGAGACACAGATATTGAGAGAGACTCAGAAGATAGCCGAGGAGAGAACACACAGAGAG	77664
Qy	1455	TGGAAGAAGACTCTGAGAAAAACACAGACACAAAGATGGAAGAGAGAGTATCAGGGTGA	1514
Db	77663	TGGAAGAGACTCTGAAAAAACACAGACACAAAGATGGAAGAGAGAGTATCAGGGTGA	77604
Qy	1515	ACACACAGTGGTGGAAATGACAAAATGCAGAGAGAAAGCAAGCAATCCAGGCGCAAGA	1574
Db	77603	ACACACAGTGGTGGAAATGACAAAATGCAGAGAGAAAGCAAGCAATCCAGGCGCAAGA	77544
Qy	1575	ATAGTGACCCAGAGTTGGTGAGAGCCAGATCCTTAAGCTGGGGGAGCAGGGAAGGGG	1634
Db	77543	ATAGTGACCCAGAGTTGGTGAGAGCCAGATCCTTAAGCTGGGGGAGCAGGGAAGGGG	77484
Qy	1635	CTGGCTTGGCTTCCGAGAGCCCTCCCCATTTCTCCGGGCGAGGAGGTAGGAGTGACAT	1694
Db	77483	CTGGCTTGGCTTCCGAGAGCCCTCCCCATTTCTCCGGGCGAGGAGGTAGGAGTGACAT	77424
Qy	1695	TCCGGACTGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCAGGAGAGCTTATTGC	1754
Db	77423	TCCGGACTGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCAGGAGAGCTTATTGC	77364
Qy	1755	TAAGGCCGATAGGCACCTCATTTGCCCGGAATGTGCCCGAGGACAGTGGGTGTTAT	1814
Db	77363	TAAGGCCGATAGGCACCTCATTTGCCCGGAATGTGCCCGAGGAGAGTGGGTGTTAT	77304
Qy	1815	AATCTCAGGCCGTGCCAGAGCCACAGGAGGAGCAGTGGCCAGGAAGGCA CAGGCTTGA	1874
Db	77303	AATCTCAGGCCGTGCCAGAGCCACAGGAGGAGGAGTGGCCAGGAAGGCA CAGGCTTGA	77244
Qy	1875	GAAGTCTGGGCTGAGCTGGAGCAAAATCCCCACCCCTACTCTGGGGACAGGCAAGT	1934



Db	77243	GAAGTCTGCGCTGAGCTGGAGCAAAATCCCCACCCCTACCTGGGGACAGGGCAAGT	77184
Qy	1935	GAGACCTGGTGGAGGTGGCTCAGCAGGAGGAAAGAGAGTGTCTGTGGCTCTGCAAC	1994
Db	77183	GAGACCTGGTGGAGGTGGCTCAGCAGGAGGAAAGAGAGTGTCTGTGGCTCTGCAAC	77124
Qy	1995	CACATCTTTCTCTGTCCCTCTCTTGGCCCTGTCTGGAGCTGTAGACTCTCTATCTCTGA	2054
Db	77123	CACATCTTTCTCTGTCCCTCTCTTGGCCCTGTCTGGAGCTGTAGACTCTCTATCTCTGA	77064
Qy	2055	ATTCTATAGTCCCTGGTCTCAGCCAGTCCGATGTGGCCCGTCTCTGTGGTTCCTCT	2114
Db	77063	ATTCTATAGTCCCTGGTCTCAGCCAGTCTGATGTGGCCCGTCTCTGTGGTTCCTTT	77004
Qy	2115	CTACCTGGGAAATAAGGTAGGGAGGAGGGAGTGGGTAAAGGCTCCCCGGATCGC	2174
Db	77003	CTACCTGGGAAATAAGGTAGGGAGGAGGGAGTGGGTAAAGGCTCCCCGGATGGC	76944
Qy	2175	CTGGGCTCCCAACCTCTGACATTCGCCATCCAGGTGACGCGCCATGGCTACAGCAAG	2234
Db	76943	CTGGGCTCCCAACCTCTGACATTCGCCATCCAGGTGACGCGCCATGGCTACAGCAAG	76884
Qy	2235	ACCCCTCGATGTGGTGTCTGTCTCTCTATCACAGCCTTGCTCTCTGGGGGTACAGG	2294
Db	76883	ACCCCTCGATGTGGTGTCTGTCTCTCTATCACAGCCTTGCTCTCTGGGGGTACAGG	76824
Qy	2295	TAACCAGAACTCTGGGGTGGAGGGTGTGGGATTTGGAGGACTGTCTCTCGCGCACTAG	2354
Db	76823	TAACCAGAACTCTGGGGTGGAGGGTGTGGGATTTGGAGGACTGTCTCTCGCGCACTAG	76764
Qy	2355	AGCGCTGTCCCTGGGAACTGTGTGAGCTGGGCACTGCTCGGACCGGGTGAATGT	2414
Db	76763	AGCGCTGTCCCTGGGAACTGTGTGAGCTGGGCACTGCTCGGACCGGGTGAATGT	76704
Qy	2415	GAGTCTCTGTCTGACTGTGGTGTGGATCGTATGTGGCCCTGTGACTGCCACGGTGT	2474
Db	76703	GAGTCTCTGTCTGACTGTGGTGTGGATCGTATGTGGCCCTGTGACTGCCACGGTGT	76644
Qy	2475	GTGTGGGGAGGGGATGCTTTTCCCATTCAGTGACTGTGGCGAGGTGGCACTGAC	2534
Db	76643	GTGTGGGGAGGGGATGCTTTTCCCATTCAGTGACTGTGGCGAGGTGGCACTGAC	76584
Qy	2535	CCTTTGAGGCTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	2594
Db	76583	CCTTTGAGGCTGTGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT	76524
Qy	2595	CACAGCTGTGTGGTGAATGATGACTAGCACTGGGGGTGTTCACTGTGTGTGGTCTGT	2654
Db	76523	CACAGCTGTGTGGTGAATGATGACTAGCACTGGGGGTGTTCACTGTGTGTGGTCTGT	76464
Qy	2655	GTGTGACTTGGCAATCTATGACTCAGGTATCTCAGTATCTCAGTATCTCAGTATCTC	2714
Db	76463	GTGTGACTTGGCAATCTATGACTCAGGTATCTCAGTATCTCAGTATCTCAGTATCTC	76404
Qy	2715	GGATTGGTGCACAAAGTGTTCATCACCATGAAAGCTGTGACTGTGTGTGTGTGTGCA	2774
Db	76403	GGATTGGTGCACAAAGTGTTCATCACCATGAAAGCTGTGACTGTGTGTGTGTGTGCA	76344
Qy	2775	GGCGATTATGTGATTGTGGCTGAGTGTGAGTGTATGATGCTCCGATTTTGTGACCGGTG	2834
Db	76343	GGCGATTATGTGATTGTGGCTGAGTGTGAGTGTATGATGCTCCGATTTTGTGACCGGTG	76284
Qy	2835	ACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTG	2894
Db	76283	ACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTGTGTGTGTGTGTG	76224
Qy	2895	TGTAATGCTACTGTATGTGTGATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	2954
Db	76223	TGTAATGCTACTGTATGTGTGATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	76166
Qy	2955	GGAGGGATAGGGGTGAGGGGTAGCTATCTCTGGGAGATGGGTGCGAGTGTGACTGCTT	3014
Db	76165	GGAGGGATAGGGGTGAGGGGTAGCTATCTCTGGGAGATGGGTGCGAGTGTGACTGCTT	76106
Qy	3015	GCAGTGTGCTGTGTGTGCAAGAGTATGTGGCAGTCTGAAATCTGTGTGCAACACCGC	3074
Db	76105	GCAGGGTGTGCTGTGTGCAAGAGTATGTGGCAGTCTGAAATCTGTGTGCAACACCGC	76046
Qy	3075	ATCTGTGCTGGTGGCACTCAGACACTGTGTGATGAGGGTGTGCGATCCCGTATGGCTGCCCG	3134
Db	76045	ATCTGTGCTGGTGGCACTCAGACACTGTGTGATGAGGGTGTGCGATCCCACTAGCTGCCCG	75986
Qy	3135	GAGCGTGTACCTGGGAGACAGAGCTGTATGTTAGCTGTACCTGTGGAGGCAACATGGGC	3194
Db	75985	GAGCGTGTACCTGGGAGACAGAGCTGTATGTTAGCTGTACCTGTGGAGGCAACATGTAC	75926
Qy	3195	GTGTCTGCAGAACTGCGCGTGTGCTTGGCTGTACTGTGTGTGTGTGTGTGTGTGTGTGG	3254
Db	75925	GTGTCTGCAGAACTGCGCGTGTGCTTGGCTGTACTGTGTGTGTGTGTGTGTGTGTGTGG	75866
Qy	3255	GTGAGTTCTGTAATGATGTGTGCTCAGGGCCATCAGCAGGGTAAAGAACAGAGCCCGGC	3314
Db	75865	GTGAGTTCTGTAATGATGTGTGCTCAGGGCCATCAGCAGGGTAAAGAACAGAGCCCGGC	75806
Qy	3315	GCGGTGCTCACGCTGTAAATCCAGCCCTTTGGGAGGCCGAGGACGCGGATCACCTGA	3374
Db	75805	GCGGTGCTCACGCTGTAAATCCAGCCCTTTGGGAGGCCGAGGACGCGGATCACCTGA	75746
Qy	3375	GGTGGGAGATCGAGGCCAGCTGTACCAACATGTGAGAAACCCCGTCTCTACTAAATAACA	3434
Db	75745	GGTGGGAGATCGAGGCCAGCTGTACCAACATGTGAGAAACCCCGTCTCTACTAAATAACA	75686
Qy	3435	AAAAATTTAGTGTGTGTGGTGGCGCTGTATCCAGACTCTCGGGAGACTGGGGCA	3494
Db	75685	AAAAATTTAGTGTGTGTGGTGGCGCTGTATCCAGACTCTCGGGAGACTGGGGCA	75626
Qy	3495	GAATAATTCGTTGAAACCCGGGAGGTGGAGTGTGGGTGAGCCGAGATCGGCCATTCAC	3554
Db	75625	GAATAATTCGTTGAAACCCGGGAGGTGGAGTGTGGGTGAGCCGAGATCGGCCATTCAC	75566
Qy	3555	TCAGCCTCGGGCAACAGAGCGGAACTCCGTCTCGAAAGAAAAAAGAAAAAAGGG	3614
Db	75565	TCAGCCTCGGGCAACAGAGCGGAACTCCGTCTCGAAAGAAAAAAGAAAAAAGGG	75506
Qy	3615	TAAGAACCTGTAATGGGACGGGAGGACTGATGATGAGTGGGCGATGATGTAGTCTG	3674
Db	75505	TAAGAACCTGTAATGGGACGGGAGGACTGATGATGAGTGGGCGATGATGTAGTCTG	75446
Qy	3675	TAGTCTGTGTGTGAGAGGAGGATGACAGGATTCAGAAAGGATGTTTTCACTCTAGA	3734
Db	75445	TAGTCTGTGTGTGAGAGGAGGATGACAGGATTCAGAAAGGATGTTTTCACTCTAGA	75386
Qy	3735	ATTGAGAAACCTAGGCTGTCTCTTCCCTCCATGTGGCCCTTAAGCTGAGCCCTTCTTT	3794
Db	75385	ATTGAGAAACCTAGGCTGTCTCTTCCCTCCATGTGGCCCTTAAGCTGAGCCCTTCTTT	75326
Qy	3795	CCTGGTCTCTTTCGGAACCTAGCTCCGCCCATGAGCTGTGACCCCACTCTCTTCTCT	3854
Db	75325	CCTGGTCTCTTTCGGAACCTAGCTCCGCCCATGAGCTGTGACCCCACTCTCTTCTCT	75266
Qy	3855	CAACACGCGCTTAGGCGAGACTCTAGTGGACCCCGCTTAAGGCCACACCCCTTTGGGCC	3914
Db	75265	CAACACGCGCTTAGGCGAGACTCTAGTGGACCCCGCTTAAGGCCACACCCCTTTGGGCC	75206
Qy	3915	AGGCTCCACCCCTATTCTGTGGGTACTTTCTAGAACCCCTTCAAAAGTCAGAGCTTTT	3974
Db	75205	AGGCTCCACCCCTATTCTGTGGGTACTTTCTAGAACCCCTTCAAAAGTCAGAGCTTTT	75149
Qy	3975	TTTTTTTTTTTTTTGGAGAGTCTGTCTCTCTCCAGGCTGAGTGTGAGTGTGAGTGTGAT	4034
Db	75148	TTTTTTTTTTTTTTGGAGAGTCTGTCTCTCTCCAGGCTGAGTGTGAGTGTGAGTGTGAT	75089
Qy	4035	CTGGGCTCACTGCAACCTCTGCGCTCCAGGTTCAAGTGTCAAGTGTCAAGTGTCAAGTGT	4094
Db	75088	CTGGGCTCACTGCAACCTCTGCGCTCCAGGTTCAAGTGTCAAGTGTCAAGTGTCAAGTGT	75029







Db	226	GCCAGAGCCAGAGAGGCGGTGGCCAGGAAGTGCAGGCTTGAGAAATCCGTGGCTG	167
QY	1889	AGCTGGGAGCAAAATCCCCCCTACCTCGGGGACAGGCGCAAGTCAGACCTGGTGAGG	1948
Db	166	AATGGNAGCAATCCCCCCTACCTCGGGGACAGGCGCAAGTCAGACCTGGGAGG	107
QY	1949	GTGGCTCAGCAGGAGGAGGAGAGTGTCGTGCGCTCTGCAACCCACATCTTCTCTG	2008
Db	106	GTGGCTCAGCAGGAGGAGGAGTGTCGTGCACTCTGCAACCCACATCTTCTCTG	47
QY	2009	TC--CCCTCTTGGCTGTCTGAGGCTGCTAGACTCTATCTTCT	2052
Db	46	TCTTCCCTCTCGCCCGCTGGAGGTGCTAGACTCTATCTTCT	1

RESULT 10

AC137591/c

LOCUS

DEFINITION

AC137591

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

32799 bp DNA linear PRI 05-DEC-2002

Homo sapiens chromosome X clone XX-B6cos map Xp22-PAR, complete sequence.

AC137591.1 GI:25446705

HTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 32799)

Rump,A., Rosenthal,A., Drescher,B. and Schattevoy,R.

Direct Submission

Submitted (07-MAY-1997) Genome Analysis, Institute Of Molecular Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

2 (bases 1 to 32799)

Lagemann,D. and Platzter,M.

Direct Submission

Submitted (26-NOV-2002) Genome Analysis, Institute Of Molecular Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany

3 (bases 1 to 32799)

Lagemann,D. and Platzter,M.

Direct Submission

Submitted (05-DEC-2002) Genome Analysis, Institute Of Molecular Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany

-----

Center: Insitute of Molecular Biotechnology

Center code: IMB

Web site: <http://genome.imb-jena.de/>

Contact: [gscj-submit@genome.imb-jena.de](mailto:gscj-submit@genome.imb-jena.de)

-----

Project Information

Center project name: B6

Center clone name: XX-540F24

-----

Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 32765 bases at least Q40

Consensus quality: 32770 bases at least Q30

Consensus quality: 32770 bases at least Q20

Quality coverage:i3.12x

-----

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one pUC18 subclone.

-----

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than 1 error in 10,000 bp.









Db	121698	GGAGTTGGAGACCGACTGGGCCAACATGGTGAACCCCTGCTCTCTACT-----A	121651
Qy	867	AAAAAAAAAAAAATCAGCCGGGTGTCGTGGGGCACACCTGTGAATCCAGCACTATGCTGGAGG	926
Db	121650	AAAAATTCAAAAATTAGCTGGGCCCTGGTGCSCGGGACCTGTAGTCCAGCTACTCAGGAGG	121591
Qy	927	CTGAGGCACGAGAGTCACTTGAAACCCCTGGAGCGGAGGTTGCAGTGGGCCGAGATCAAT	986
Db	121590	CTGAGGCAGGAGAAATCGCTGGAACCTGGGAGGCAGAGGTTGCAATGAGCCGAGATTGC	121531
Qy	987	CACCGCCCTCCAGCGCTGGGCCACAGACGACAGACTCTGTCTCAAAATAAATAAACA	1046
Db	121530	CACGTGCACTCAGCGCTGGGTGACAGAAATGAGACTCCATCTCAAAAATAAATAATAGT	121471
Qy	1047	CGAACACAGACGATTTGTTGTACCTTAGTTATA	1077
Db	121470	AAATAAACTAGACAGACTCTGTTGACTAACA	121440

RESULT 14

[illegible]

3': RP11-624D20 (UWGC:bc0494) AC068222

-----

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

-----

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

Overlapping Sequences:  
5': RP11-349E16 (UWGC:bc0646) AC098481



[illegible]



Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.

# Direct Submission

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

2 (bases 1 to 227137)  
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
Direct Submission  
Submitted (23-OCT-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 227137)  
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,  
Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C.  
and Haugen, E.D.  
Direct Submission  
Submitted (15-JUN-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Jun 15, 2002 this sequence version replaced gi:16328287.

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: [uwgchgs@u.washington.edu](mailto:uwgchgs@u.washington.edu)

Project Information

Center project name: Chr-3

Center clone name: RP11-349E16 (bc0646)

Summary Statistics

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator ET; 76% of reads

Chemistry: Dye-terminator Big Dye; 24% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 227062 bases at least Q40

Consensus quality: 227131 bases at least Q30

Consensus quality: 227137 bases at least Q20

Insert size: 227137; sum-of-contigs

Quality coverage: 8.0x in Q20 bases; sum-of-contigs

## Overlapping Sequences:

5': RP11-728B21 (UWGC:bc0540) AC114876, 123790-bp overlap

3': RP11-286L5 (UWGC:bc0344) AC103559

## Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:

all regions were either double-stranded or sequenced with an

alternate chemistry or covered by high quality data (i.e., Phred

quality > 30); an attempt was made to resolve all sequencing

problems, such as compressions and repeats; all regions were

covered by at least one plasmid subclone or more than one M13

subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically digested sequence consists of both insert and  
vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-800 bp)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered  
fragments are separated by dashed lines.

NsiI

EcoRI

HindIII

SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt  
-----  
SeqDerMap FngPrnt SeqDerMap FngPrnt SeqDerMap FngPrnt

22527	22843	8696	8633	3211	3196
7364	7368	6	<800	6382	6483
1250	1186	4641	4671	512	<800
1646	1681	924	937	449	<800
5862	5809	2669	2702	1398	1327
4415	4376	2286	2141	1649	1684
5203	5251	2356	2346	2341	2357
2339	2380	9092	9731	971	998
4104	4145	13999	13681	7003	7067
2067	2088	6660	6723	3223	3196
1413	1416	286	<800	1336	1327
8832	8730	8529	8633	2278	2251
1451	1416	11874	11716	5086	5023
3108	3003	10800	10477	587	<800
1155	1088	2842	2850	4763	4697
10132	9934	5654	5704	192	<800
7178	7368	8640	8633	16514	16792
2392	2380	13221	12924	3041	3013
4819	4786	555	<800	1803	1813
1188	1186	5455	5489	1123	1071
2666	2689	3694	3751	3332	3382
1188	1186	4919	4887	1092	1071
431	<800	951	937	4492	4528
732	750	334	<800	7571	7668
1680	1681	3403	3506	4767	4697
1855	1877	3587	3633	817	832
1537	1490	2819	2850	1340	1327
1361	1350	3838	3909	1004	998
5212	5251	5684	5704	891	922
24	<800	373	<800	10593	10357
1608	1681	63	<800	2970	2846
1731	1877	203	<800	9418	9439
1439	1416	7896	7870	1527	1526
327	<800	332	<800	2795	2775
5312	5251	407	<800	1052	1071
3221	3262	3904	3909	305	<800

Db	91802	GCTCATCTAATGAAGAAATTATGACAGGCTGTATAAACTAGGCAGGGGATGGGCATGA	91861
QY	746	TGGCTCATGCTCTAATTCACGACATTTGGGAGGCAGAGGTGGGTGAATCACTCTGAGGCC	805
Db	91862	TGGCTCACGCTGTAAATCCAGACATTTGGGAGGCCAAGGCCAGTGGATCACTTGAGTTC	91921
QY	806	AGGAGTTTGAGACCGCTGGCCAACATGGTGAAACTCTGTCTTTACTAAAAAATAAAAA	865
Db	91922	CGGAGTTGAGACCGATCTGGCCACGTTGTAACCCCATCTTACT	91969
QY	866	AAAAAATAAAAAATCAGCCGGGTGTTCGTGGGCACACCTGTAAATCCACGCTATGCTGGAG	925
Db	91970	AAAAATACAAAAATTAGCCAGGCTTAGTGGTGTGTGCTGTGTAGTCCAGCTACTCGGNG	92029
QY	926	GCTGAGGCACGAGAGTCACTTGAAACCTCGAGGCGGAGGTTGCAGTGGGCCGAGATCACA	985
Db	92030	TCGTGAGCATGAGAAATCACTTGAACCTGGGAGGTGGAGGTTGCAGTGAGCCCAAGATTACG	92089
QY	986	TCACCGCCCTCCAGCCTGGGCGACAGACAAGACTCTGTCTCAAAATAAATAATAACAA	1045
Db	92090	CCACTGTACTCCAGCCTGGGTGACAGACAAGACTCTGTCTCAAAAAATAAATAATAA	92149
QY	1046	ACGAA 1050	
Db	92150	AATAAA 92154	

Search completed: July 4, 2005, 09:55:23  
Job time : 14430.2 secs

[illegible]

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 1759.59 Seconds

(without alignments)  
16821.327 Million cell updates/sec

Title: US-09-936-271c-13\_COPY\_1\_5000

Perfect score: 5000

Sequence: 1 gggccagagtggaagcaag.....tgatccacacagtggcttc 5000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1980s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5000	100.0	11570	3 AAA95905	AAA95905 Human KLK
2	5000	100.0	11570	12 ADK52482	Adk52482 Human kal
3	5000	100.0	11570	13 ADR72623	Adr72623 Human ren
4	5000	100.0	11570	13 ADR72875	Adr72875 Human ova
5	472	9.4	525	12 ACH80345	Ach80345 Human gen
6	440.2	8.8	41454	11 ACN44942	Acn44942 Human gen
7	424.4	8.5	84607	2 AAX90847	Aax90847 Human PAC
8	419	8.4	91352	12 ADN94799	Adn94799 DNA encod
9	418.4	8.4	115935	13 ADR52890	Adr52890 Drug ther
10	417.6	8.4	26371	11 ACN44814	Acn44814 Human gen
11	417.6	8.4	91760	11 ACN44410	Acn44410 Human gen
12	417.4	8.3	216215	10 ADF69167	Adf69167 Human MP5
13	416.6	8.3	139904	6 ABK83562	Abk83562 Human cDN
14	416.2	8.3	79652	10 ADG70184_3	Continuation (4 of
15	415.2	8.3	17397	4 AAS36445	Aas36445 Human car
16	415.2	8.3	17397	10 ADE47139	Ade47139 Human car
17	415.2	8.3	17397	13 ADJ08557	Adj08557 Human car
18	415.2	8.3	19334	4 AAS36443	Aas36443 Human car
19	415.2	8.3	19334	10 ADE47137	Ade47137 Human car
20	415.2	8.3	19334	13 ADJ08555	Adj08555 Human car

21	415.2	8.3	19345	4 AAS36444	Aas36444 Human car
22	415.2	8.3	19345	10 ADK47138	Adk47138 Human car
23	415.2	8.3	19345	13 ADJ08556	Adj08556 Human car
24	415.2	8.3	29001	12 ADO85575	Ado85575 Human zin
25	414.8	8.3	12970	4 AAL05001	Aal05001 Human rep
26	414.8	8.3	12970	4 ABL97894	Ab197894 Human tes
27	414.6	8.3	74037	6 ABK94412	Abk94412 DNA encod
28	414.6	8.3	111084	12 ADQ18808	Adq18808 Human sof
29	414.2	8.3	212231	11 ACN44598	Acn44598 Human gen
30	410.4	8.2	10445	4 AAL03692	Aal03692 Human rep
31	409.6	8.2	50000	9 ADB16926	Adb16926 Human DYX
32	408.2	8.2	240000	8 ACD13446	Acd13446 Human DNA
33	407.6	8.2	196686	11 ACN44170	Acn44170 Human gen
34	407.6	8.2	196686	12 ADQ07738	Adq07738 Nucleotid
35	406.2	8.1	12970	4 AAL05001	Aal05001 Human rep
36	406.2	8.1	12970	4 ABL97894	Ab197894 Human tes
37	406	8.1	32224	4 AAK89986	Aak89986 Human dig
38	405.8	8.1	4433	4 AAL04142	Aal04142 Human rep
39	405.8	8.1	4433	5 ABA18400	Aba18400 Human ner
40	405.8	8.1	4433	5 AAS40516	Aas40516 DNA encod
41	405.8	8.1	4433	11 ADJ09722	Adj09722 Human pro
42	405.8	8.1	126001	12 ADH77123	Adh77123 Human PAZ
43	405	8.1	32767	8 AAD56075	Aad56075 Human SNL
44	405	8.1	32767	9 ADA02437	Ada02437 Human SNL
45	405	8.1	32767	10 ADB72176	Adb72176 Human SNL

## ALIGNMENTS

RESULT 1  
AAA95905  
ID AAA95905 standard; DNA; 11570 BP.  
XX  
AC AAA95905;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KLK-L2 gene.  
XX  
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
KW kallikrein-like protein; serine protease; cytosolic; cancer;  
KW prostrate cancer; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
XX 09-MAR-2000; 2000WO-CA000258.  
XX  
PR 11-MAR-1999; 99US-0124260P.  
PR 01-APR-1999; 99US-0127386P.  
PR 21-JUL-1999; 99US-0144919P.  
XX  
PA (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
PI Yousef GM, Diamandis EP;  
XX  
XX WPI; 2000-587440/55.  
DR P-PSDB; AAB21296.  
XX  
PT New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
XX protein mediated disorders, especially cancer.  
PS Claim 1; Page 143-149; 184pp; English.  
XX  
XX The present sequence is the coding sequence of the human KLK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyse the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding



1981 GTGGGCTCTGACCAACATCTTTCTCTGTCCCTCCCTTGTGACCTGTCTGGAGGCTGTCTAGA 2040  
1982 |||||  
2041 CTCCTATCTCTGAAATCTATAGTCCCTGAGCTCTCAGCGAGTCCAGATGGTGCCTGTC 2100  
2042 |||||  
2043 CTCCTATCTCTGAAATCTATAGTCCCTGAGCTCTCAGCGAGTCCAGATGGTGCCTGTC 2100  
2044 |||||  
2101 CTTGTGGTTCCTCTCTACCTGGGGAATAAGTATAGGAGGAGGAGGAGTGGGTTAAGG 2160  
2102 |||||  
2103 CTTGTGGTTCCTCTCTACCTGGGGAATAAGTATAGGAGGAGGAGGAGTGGGTTAAGG 2160  
2104 |||||  
2161 GCTCCCGGATCGCTGGGCTCCCAACCTCTCAGCAATCTCCCAATCCAGGTGACAGCGGCC 2220  
2162 |||||  
2163 GCTCCCGGATCGCTGGGCTCCCAACCTCTCAGCAATCTCCCAATCCAGGTGACAGCGGCC 2220  
2164 |||||  
2221 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTGCTCTGTATCAGAGCTTGTCTT 2280  
2222 |||||  
2223 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGTGCTCTGTATCAGAGCTTGTCTT 2280  
2224 |||||  
2281 CTGGGGTTCAGAGTAACAGAACTCTGGGGTGGGAGGTTGTGGATTGGGAGGACTGT 2340  
2282 |||||  
2283 CTGGGGTTCAGAGTAACAGAACTCTGGGGTGGGAGGTTGTGGATTGGGAGGACTGT 2340  
2284 |||||  
2341 CTCTGGGCACTAGAGCGCTGTCCCTGGGAACTGTGTGAGCTGGGATGACTCCGG 2400  
2342 |||||  
2343 CTCTGGGCACTAGAGCGCTGTCCCTGGGAACTGTGTGAGCTGGGATGACTCCGG 2400  
2344 |||||  
2401 GACCGGTGAAATGAGTCTCTGTCTGTACTTGTGGTTGTGCGATCGTATGTGCGCCCTGT 2460  
2402 |||||  
2403 GACCGGTGAAATGAGTCTCTGTCTGTACTTGTGGTTGTGCGATCGTATGTGCGCCCTGT 2460  
2404 |||||  
2461 GACTGCCACGCTGTGTGGGGAGGGGATGCTCTTCCATATCAGGTGACTGTGCGG 2520  
2462 |||||  
2463 GACTGCCACGCTGTGTGGGGAGGGGATGCTCTTCCATATCAGGTGACTGTGCGG 2520  
2464 |||||  
2521 CAGGTGGCACTGACCTTTGAGGCTGTGTGTGCTGTTTGTGATTGTGTGCAATTAAAG 2580  
2522 |||||  
2523 CAGGTGGCACTGACCTTTGAGGCTGTGTGTGCTGTTTGTGATTGTGTGCAATTAAAG 2580  
2524 |||||  
2581 ATTGTGTGTGCTCCACAGCTGTGTGGGTGAATGCATGTAGCACTGGGGGTGTTCACCTGT 2640  
2582 |||||  
2583 ATTGTGTGTGCTCCACAGCTGTGTGGGTGAATGCATGTAGCACTGGGGGTGTTCACCTGT 2640  
2584 |||||  
2641 GTGTTTGGCTGTGTGTGCTGCTGGCAATGTATGACTCAGGTATCTGCAGTTCCTG 2700  
2642 |||||  
2643 GTGTTTGGCTGTGTGTGCTGCTGGCAATGTATGACTCAGGTATCTGCAGTTCCTG 2700  
2644 |||||  
2701 TCCCTGAGTCCCGGATTCGTGCAACAAAGTGTTCATCACCATGGAAAGCTGTGACT 2760  
2702 |||||  
2703 TCCCTGAGTCCCGGATTCGTGCAACAAAGTGTTCATCACCATGGAAAGCTGTGACT 2760  
2704 |||||  
2761 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGAGCTTATGATGCCGTA 2820  
2762 |||||  
2763 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGAGCTTATGATGCCGTA 2820  
2764 |||||  
2821 TTTGTGACCGGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880  
2822 |||||  
2823 TTTGTGACCGGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG 2880  
2824 |||||  
2881 TCTGTGTGAGCGCGTGTAAATGCTATCTGTATGTGTGATGTGTGAGTGTGTGTGAGT 2940  
2882 |||||  
2883 TCTGTGTGAGCGCGTGTAAATGCTATCTGTATGTGTGATGTGTGAGTGTGTGTGAGT 2940  
2884 |||||  
2941 TTTCTGTCTCTCCCTGGAGGATAGAGGTGAGGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
2942 |||||  
2943 TTTCTGTCTCTCCCTGGAGGATAGAGGTGAGGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
2944 |||||  
3001 CAGGTGACTGACTGTGCTGTGCTGTGTGCAAGAGATGTGTGAGTGTGCAACATC 3060  
3002 |||||  
3003 CAGGTGACTGACTGTGCTGTGCTGTGTGCAAGAGATGTGTGAGTGTGCAACATC 3060  
3004 |||||  
3061 TGTGCACACCGGATCTGTGCTGGCACTGAGCACTGTGGAATGAGGGTGTGCAATCCC 3120  
3062 |||||  
3063 TGTGTCTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGTGCTCTCAACTCCCAACC 4200  
3064 |||||

3061 TGTGCACACAGGCACTCTGTGCTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC 3120  
3062 |||||  
3121 GCTAGGCTGCCCGGAGCGTGTGTACTGTGAGACAGAGCTGTATGTTAGCTGCACTGTG 3180  
3122 |||||  
3123 GCTAGGCTGCCCGGAGCGTGTGTACTGTGAGACAGAGCTGTATGTTAGCTGCACTGTG 3180  
3124 |||||  
3181 GAGCAACATGGGCGTGTCTGCAAGAACTGCTGCTGCTTGGCTGTACTGCTGCTTGTGC 3240  
3182 |||||  
3183 GAGCAACATGGGCGTGTCTGCAAGAACTGCTGCTGCTTGGCTGTACTGCTGCTTGTGC 3240  
3184 |||||  
3241 GCGTGGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTCGCCAGGCGCATCAGCAAGGTA 3300  
3242 |||||  
3243 GCGTGGTCTTGTGGGTGAGTTCGTGAATGATGTGTGTCGCCAGGCGCATCAGCAAGGTA 3300  
3244 |||||  
3301 GAACAGGCGCGGCGGCTGCTCAGCCTGTAACTCCAGCCCTTTGGGAGGCGGAGCA 3360  
3302 |||||  
3303 GAACAGGCGCGGCGGCTGCTCAGCCTGTAACTCCAGCCCTTTGGGAGGCGGAGCA 3360  
3304 |||||  
3361 GCGGATCACCTGAGGTCCGGAGATCGAGGCCAGCCTGACCAATGAGAGAACCCCTCT 3420  
3362 |||||  
3363 GCGGATCACCTGAGGTCCGGAGATCGAGGCCAGCCTGACCAATGAGAGAACCCCTCT 3420  
3364 |||||  
3421 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGCGCTGCTGTATATCCAGCTACTC 3480  
3422 |||||  
3423 CTACTAAAAATACAAAAATTTAGCTGTGTGTGGCGCTGCTGTATATCCAGCTACTC 3480  
3424 |||||  
3481 GCGAGACTGGGCGAGAAAAATTCGCTTGAACCCGGGAGGTGGAGTTGCGGTGAGCCGAGA 3540  
3482 |||||  
3483 GCGAGACTGGGCGAGAAAAATTCGCTTGAACCCGGGAGGTGGAGTTGCGGTGAGCCGAGA 3540  
3484 |||||  
3541 TCGGCGCATTTGCACTCCAGCCTTGGGCAACAGAGCGGAACTCCGCTCTCGAAGAAAAA 3600  
3542 |||||  
3543 TCGGCGCATTTGCACTCCAGCCTTGGGCAACAGAGCGGAACTCCGCTCTCGAAGAAAAA 3600  
3544 |||||  
3601 GAAAAAAGGAGGTGAAGAACTGATGAAATGGGCAACGGGAGGACTGATGATGGAGTGGGC 3660  
3602 |||||  
3603 GAAAAAAGGAGGTGAAGAACTGATGAAATGGGCAACGGGAGGACTGATGATGGAGTGGGC 3660  
3604 |||||  
3661 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGAT 3720  
3662 |||||  
3663 ATGCATGTAGTCTGTAGTCTGTGTGAGAGGAGGAGATTGACAGGATTGAGAAGGAT 3720  
3664 |||||  
3721 GTTTTTCATCTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGCGCCCTTAAG 3780  
3722 |||||  
3723 GTTTTTCATCTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTCCATGTGCGCCCTTAAG 3780  
3724 |||||  
3781 CTGAGCCCTTCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 3840  
3782 |||||  
3783 CTGAGCCCTTCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 3840  
3784 |||||  
3841 CCACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3900  
3842 |||||  
3843 CCACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3900  
3844 |||||  
3901 CACCCCTTGTGGGCGAGGCTCCACCCCTTATCTGTGGGTACCTTCTAGAACCCCTTCAA 3960  
3902 |||||  
3903 CACCCCTTGTGGGCGAGGCTCCACCCCTTATCTGTGGGTACCTTCTAGAACCCCTTCAA 3960  
3904 |||||  
3961 AGTCAGAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4020  
3962 |||||  
3963 AGTCAGAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4020  
3964 |||||  
4021 TGCAGTGGCTGTATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080  
4022 |||||  
4023 TGCAGTGGCTGTATCTCGGCTCACTGCAACCTCTGCTCCAGGTTCAAGTGAATTCGT 4080  
4024 |||||  
4081 GCCTTCCACCTCCTGAGTGTGGATTACAGGTGCGCGCCACACGCTGCTGCTAAATTTT 4140  
4082 |||||  
4083 GCCTTCCACCTCCTGAGTGTGGATTACAGGTGCGCGCCACACGCTGCTGCTAAATTTT 4140  
4084 |||||  
4141 GTGTCTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGTGCTCTCAACTCCCAACC 4200  
4142 |||||  
4143 GTGTCTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGTGCTCTCAACTCCCAACC 4200  
4144 |||||

QY 4201 TCAGGTGATCGGCCACCTCGGCTCCAGAGTCTGGGGTTACAGCGTGGAGCCACGC 4260  
DB TCAGGTGATCGGCCACCTCGGCTCCAGAGTCTGGGGTTACAGCGTGGAGCCACGC 4260  
QY 4261 CCCAGGCCAAAGTCAGAGCTCTTTATAGAGACTCTAAACATGTAAACCTGTGGCC 4320  
DB CCCAGGCCAAAGTCAGAGCTCTTTATAGAGACTCTAAACATGTAAACCTGTGGCC 4320  
QY 4321 CTAACCTAAGTCAATTCCTCAACCCCTCTCGCTCCAGCCCTGGTCTCCCTCCCATCCCCAG 4380  
DB CTAACCTAAGTCAATTCCTCAACCCCTCTCGCTCCAGCCCTGGTCTCCCTCCCATCCCCAG 4380  
QY 4381 TGACCCCACTCTTTGAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440  
DB TGACCCCACTCTTTGAGACAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440  
QY 4441 GCTCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAT 4500  
DB GCTCAGCCCCACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCCTGTACCCAAT 4500  
QY 4501 TTTACCTCAGATGAGTCTAGCAATTCAGGAATCTGTGAGTCCAGTTAGATCCCA 4560  
DB TTTACCTCAGATGAGTCTAGCAATTCAGGAATCTGTGAGTCCAGTTAGATCCCA 4560  
QY 4561 GTAAACCTACCTGAGCTGGGCTCTGTCTTGTAGCTTGGCTGGGTTGAGAGTGCCA 4620  
DB GTAAACCTACCTGAGCTGGGCTCTGTCTTGTAGCTTGGCTGGGTTGAGAGTGCCA 4620  
QY 4621 CTCTTATCTCAGGCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680  
DB CTCTTATCTCAGGCTCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4680  
QY 4681 GGCTGCTCTGAGTCTGAACCCACCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740  
DB GGCTGCTCTGAGTCTGAACCCACCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4740  
QY 4741 CAACCCATTTTCGTTCCAGAGATGTTCTCGGCAACATGATGTTCTGTGACCAAC 4800  
DB CAACCCATTTTCGTTCCAGAGATGTTCTCGGCAACATGATGTTCTGTGACCAAC 4800  
QY 4801 CCTCTAACACCGTCCCTCTGGAGCAACAGGACCTGGGAGCTGGGGCCGGGGAAGCG 4860  
DB CCTCTAACACCGTCCCTCTGGAGCAACAGGACCTGGGAGCTGGGGCCGGGGAAGCG 4860  
QY 4861 CCGGTGCGATGACAGCAGCGGCTCATCATGATGATGATGATGATGATGATGATGATGATG 4920  
DB CCGGTGCGATGACAGCAGCGGCTCATCATGATGATGATGATGATGATGATGATGATGATG 4920  
QY 4921 AGCGTGGCAGGCGCGCTGTTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980  
DB AGCGTGGCAGGCGCGCTGTTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980  
QY 4981 TGCATCCACAGTGGCTGCTC 5000  
DB TGCATCCACAGTGGCTGCTC 5000

RESULT 2

ADK52482  
ID ADK52482 standard; DNA; 11570 BP.

XX AC ADK52482;

XX DT 03-JUN-2004 (first entry)

XX DE Human kallikrein 5 encoding sequence.

XX KW kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.

XX OS Homo sapiens.

XX PN W02004021008-A2.

XX 11-MAR-2004.  
XX 28-AUG-2003; 2003WO-CA001310.  
XX 28-AUG-2002; 2002US-0407333P.  
XX (MOUN ) MOUNT SINAI HOSPITAL.  
XX (YOUS/) YOUSEF G.  
XX Diamandis EP;  
XX WPI; 2004-239232/22.  
XX P-PSDB; ADK52481.  
XX Detecting kallikrein 5 associated with cancer, useful in diagnosing,  
XX monitoring, detecting, imaging and treating breast or ovarian carcinoma,  
XX comprises comparing the detected amount of kallikrein 5 in a sample with  
XX a standard sample.  
XX Disclosure; SEQ ID NO 2; 59pp; English.  
XX The present inventin relates to detecting kallikrein 5 associated with  
XX breast or ovarian cancer in a patient comprises detecting in the sample  
XX kallikrein 5 and comparing the detected amount with an amount detected  
XX for a standard. The method is useful in detecting kallikrein 5 associated  
XX with breast or ovarian cancer in a patient. The methods and kits are  
XX useful in diagnosing, monitoring, detecting, imaging and treating breast  
XX or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for  
XX preventing and treating breast and ovarian cancer and for stimulating or  
XX enhancing antibody production or for inducing an immune response. The  
XX present sequence represents human kallikrein 5 encoding sequence.  
XX SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;  
Query Match 100.0%; Score 5000; DB 12; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGCCAGAGTGAAGCAAGAGAGAGTGGAGAGTCCCTCTGCAAGTGGTGGTGC 60  
DB 1 GGGCCAGAGTGAAGCAAGAGAGAGTGGAGAGTCCCTCTGCAAGTGGTGGTGC 60  
QY 61 TCCCTCGCTCTAAATCGGGAGAGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
DB 61 TCCCTCGCTCTAAATCGGGAGAGGGAGGAGGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 121 AAGAAAG 180  
DB 121 AAGAAAG 180  
QY 181 AAG 240  
DB 181 AAG 240  
QY 241 ACACAAATGAG 300  
DB 241 ACACAAATGAG 300  
QY 301 AAAGGGCAG 360  
DB 301 AAAGGGCAG 360  
QY 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 420  
DB 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 420  
QY 421 GATCTCGGCTCACTGCAAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 480  
DB 421 GATCTCGGCTCACTGCAAGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 480  
QY 481 CCAAGTAGCTGGAGCTACAGGGCCCGCCACTACGCGCCGGCTAAATTTTGTATTATTTA 540  
DB 481 CCAAGTAGCTGGAGCTACAGGGCCCGCCACTACGCGCCGGCTAAATTTTGTATTATTTA 540





QY 2701 TCCCTGAGGTCCTCCGGGATTTGGTGCACAAAGAGTGTCTATCAACATGGAAAGCTGTGACT 2760  
DB 2701 TCCCTGAGGTCCTCCGGGATTTGGTGCACAAAGAGTGTCTATCAACATGGAAAGCTGTGACT 2760  
QY 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGACGTTATGGATGCCCGTA 2820  
DB 2761 GTGTGCTGCTTGCAGGCGATTATGTGATTGTGGCTGAGTGTGACGTTATGGATGCCCGTA 2820  
QY 2821 TTTGTGACCGTGTGACTACCTGAAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880  
DB 2821 TTTGTGACCGTGTGACTACCTGAAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG 2880  
QY 2881 TCTGTGTAGCGCGTGTAAATGTCTATGTATGTGTGATGTGTGAGCTGTGTGTGAGT 2940  
DB 2881 TCTGTGTAGCGCGTGTAAATGTCTATGTATGTGTGATGTGTGAGCTGTGTGTGAGT 2940  
QY 2941 TTCTGTCTCTGCTCGAGGAGATAGAGGTGCAGGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
DB 2941 TTCTGTCTCTGCTCGAGGAGATAGAGGTGCAGGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
QY 3001 CAGGTGACTGACTTGCAGTGTGTGCTGTGTGCAGAAAGATGTGTGGCAGTCTGAAACATC 3060  
DB 3001 CAGGTGACTGACTTGCAGTGTGTGCTGTGTGCAGAAAGATGTGTGGCAGTCTGAAACATC 3060  
QY 3061 TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC 3120  
DB 3061 TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC 3120  
QY 3121 GCTTAGGTCGCCGGAGCGTGTACCTGAGACAGAGCTGTATGTAGCTGCACCTGTG 3180  
DB 3121 GCTTAGGTCGCCGGAGCGTGTGTACCTGAGACAGAGCTGTATGTAGCTGCACCTGTG 3180  
QY 3181 GAGGCAACATGGGCGTCTGTGCAGAACTGCGTGGTCTTGGCTGTACTGTGTGTGTC 3240  
DB 3181 GAGGCAACATGGGCGTCTGTGCAGAACTGCGTGGTCTTGGCTGTACTGTGTGTGTC 3240  
QY 3241 GCGTGTGTTCTGGGGTGAGTTGCGTAATGATGTGTGTCGAGGCGCATCAGCAAGGGTAA 3300  
DB 3241 GCGTGTGTTCTGGGGTGAGTTGCGTAATGATGTGTGTCGAGGCGCATCAGCAAGGGTAA 3300  
QY 3301 GAACAGGCGGGCGCGTGGCTCAGCCCTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360  
DB 3301 GAACAGGCGGGCGCGTGGCTCAGCCCTGTAATCCAGCCCTTTGGGAGGCGGAGCA 3360  
QY 3361 GGGGATCACTGAGGTCGGGAGATCGAGGCGAGCTGACCAACATGGAGAACCCCGTCT 3420  
DB 3361 GGGGATCACTGAGGTCGGGAGATCGAGGCGAGCTGACCAACATGGAGAACCCCGTCT 3420  
QY 3421 CTACTAAAAATACAAAAATTAGCTGTGTGGTGGCGCGTGTAAATCCAGCTACTC 3480  
DB 3421 CTACTAAAAATACAAAAATTAGCTGTGTGGTGGCGCGTGTAAATCCAGCTACTC 3480  
QY 3481 GGGAGCTGGGCGAGAAAAATCCGTTGAACCCGGAGGTGGAGTTGCGGTGAGCCGAGA 3540  
DB 3481 GGGAGCTGGGCGAGAAAAATCCGTTGAACCCGGAGGTGGAGTTGCGGTGAGCCGAGA 3540  
QY 3541 TCGCGCATTTGCATCTCAGCCTGGGCAACAAGACGAAACTCCGCTCTCGAAAGAAAAAA 3600  
DB 3541 TCGCGCATTTGCATCTCAGCCTGGGCAACAAGACGAAACTCCGCTCTCGAAAGAAAAAA 3600  
QY 3601 GAAAAAAGGTAAGAACCAAGTGAATGGCAACGGGAGCTGATGATGGAGTGGGC 3660  
DB 3601 GAAAAAAGGTAAGAACCAAGTGAATGGCAACGGGAGCTGATGATGGAGTGGGC 3660  
QY 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATGTGACAGGATTTGAGAGGCAT 3720  
DB 3661 ATGCATGTAGTCTGTAGTCTGTGTGTGAGAGGAGATGTGACAGGATTTGAGAGGCAT 3720  
QY 3721 GTTTTCACTGAGAAATTCAGAAACCTTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780  
DB 3721 GTTTTCACTGAGAAATTCAGAAACCTTAGGCTGTCTTTCCCTCCATGTGGCCCTTAAG 3780  
QY 3781 CTGAGCCCTTCTTTCTCGGTCTGCTTTTCGGAACCCCTAGCTCCGCGCATGAGCTCTGACC 3840

DB 3781 CTGAGCCCTTCTTTCTCGGTCTGCTTTTCGGAACCCCTAGCTCCGCGCATGAGCTCTGACC 3840  
QY 3841 CCACCTCCCTTTCCTCAACACAGCCCTAGGCCAGACTCTTAGTGAGACCCCGCTTAAGGCCA 3900  
DB 3841 CCACCTCCCTTTCCTCAACACAGCCCTAGGCCAGACTCTTAGTGAGACCCCGCTTAAGGCCA 3900  
QY 3901 CACCCCTTTGGGCGAGGCTCCACCCCTTATTTCTGTGGGTACCTTTAGAAACCCCTTCAA 3960  
DB 3901 CACCCCTTTGGGCGAGGCTCCACCCCTTATTTCTGTGGGTACCTTTAGAAACCCCTTCAA 3960  
QY 3961 AGTCAGAGCTTTTCTTTTGTGGAGACAGCTTTGCTCTCTCTCCAGGCTGGAG 4020  
DB 3961 AGTCAGAGCTTTTCTTTTGTGGAGACAGCTTTGCTCTCTCTCCAGGCTGGAG 4020  
QY 4021 TGCAGTGGCGTGATCTCGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGTGATTTCTGT 4080  
DB 4021 TGCAGTGGCGTGATCTCGGCTCACTGCAACCTCTGCTCTCCAGGTTCAAGTGATTTCTGT 4080  
QY 4081 GCCTCCACCTCTCTGAGTAGTGGGATTTACAGGTGCGCGGCCACACACGCTGGCTAAATTTT 4140  
DB 4081 GCCTCCACCTCTCTGAGTAGTGGGATTTACAGGTGCGCGGCCACACACGCTGGCTAAATTTT 4140  
QY 4141 GTGTCTTTTGTAGTAGACAGGGTTTCACTTGTGTGGCGAGGCTGTCTCAAACTCCCAACC 4200  
DB 4141 GTGTCTTTTGTAGTAGACAGGGTTTCACTTGTGTGGCGAGGCTGTCTCAAACTCCCAACC 4200  
QY 4201 TCAGGTGATCCGCGCCACTCGGCTCCAGAGTGTGGGGTTACAGGCGTGAGCCACCGC 4260  
DB 4201 TCAGGTGATCCGCGCCACTCGGCTCCAGAGTGTGGGGTTACAGGCGTGAGCCACCGC 4260  
QY 4261 CCCAGCCCAAGTCAAGACTCTTTATAGGAGACTCTAAACATGTAACCTGACCTCCGACC 4320  
DB 4261 CCCAGCCCAAGTCAAGACTCTTTATAGGAGACTCTAAACATGTAACCTGACCTCCGACC 4320  
QY 4321 CTAACTAAGTCAATTTCCAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380  
DB 4321 CTAACTAAGTCAATTTCCAAACCCCTTCTGCTCCAGCCCTGACCCCACTCACTGAGGCC 4380  
QY 4381 TGACCCCACTTCTTGAGACAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440  
DB 4381 TGACCCCACTTCTTGAGACAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCATCCCCAG 4440  
QY 4441 GCTCCAGCCCCACACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCTGTACCCAAT 4500  
DB 4441 GCTCCAGCCCCACACAGCTTTGGCACTACCCCTGAGCTTGTCCAGGAATCTGTACCCAAT 4500  
QY 4501 TTTACCCCTCACATGTATGTCTAGCCAAATTCAGGAAATCTGTGAGGTCCAGTTAGAGTCCA 4560  
DB 4501 TTTACCCCTCACATGTATGTCTAGCCAAATTCAGGAAATCTGTGAGGTCCAGTTAGAGTCCA 4560  
QY 4561 GTAACCCCTACCTGAGCTGGCTCTGCTTGAGCTTGGCCCTGGGCTTGAGAGGTGCCA 4620  
DB 4561 GTAACCCCTACCTGAGCTGGCTCTGCTTGAGCTTGGCCCTGGGCTTGAGAGGTGCCA 4620  
QY 4621 CTCTTATTTCTCCAGGCGCTGCCCTGCCCTCAGCATGTCCAGACACCCCACTCTAGCT 4680  
DB 4621 CTCTTATTTCTCCAGGCGCTGCCCTGCCCTCAGCATGTCCAGACACCCCACTCTAGCT 4680  
QY 4681 GGTCTGGCTCTTGTAGTCTGAAACCCCAACCCCGCCAGCCCGCCCTCTGAGCCCGCC 4740  
DB 4681 GGTCTGGCTCTTGTAGTCTGAAACCCCAACCCCGCCAGCCCGCCCTCTGAGCCCGCC 4740  
QY 4741 CAAACCATTTTCCGTTCCAGAGCATGTTCTCGCCAAATGATGTTCTCTGTGACCAAC 4800  
DB 4741 CAAACCATTTTCCGTTCCAGAGCATGTTCTCGCCAAATGATGTTCTCTGTGACCAAC 4800  
QY 4801 CCTCTAACACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGAAGACG 4860  
DB 4801 CCTCTAACACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGAAGACG 4860  
QY 4861 CCGGTCGGATGACAGAGCAGCGCATCATTAATGATCCGATCGCATATGACACCC 4920



Db 4861 CCGGTGGATGACAGCAGCGCGCATCATCATGATCGACTCGATATGACACACC 4920  
Qy 4921 AGCGTGGCAGCGCGCTGTGTCTAAGGCCCAACAGCTTACTTGGGGCGGTGTGG 4980  
Db 4921 AGCGTGGCAGCGCGCGCTGTGTCTAAGGCCCAACAGCTTACTTGGGGCGGTGTGG 4980  
Qy 4981 TGCATCCACAGTGGCTGCTC 5000  
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 3  
ADR72623  
ID ADR72623 standard; DNA; 11570 BP.  
AC ADR72623;  
XX  
XX 02-DEC-2004 (first entry)  
XX Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.  
XX kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;  
KW cancer metastasis; chemotherapy; human; serine protease;  
KW chromosome 19q13.4; KLK5; ds; gene.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 2221..11247  
FT /\*tag= b  
FT /product= "Human renal cell carcinoma-related kallikrein  
FT 5 (hK5) protein"  
FT exon 2221..2293  
FT /\*tag= a  
FT /number= 1  
FT intron 2294..4761  
FT /\*tag= c  
FT /number= 1  
FT exon 4762..5023  
FT /\*tag= d  
FT /number= 2  
FT intron 5024..5762  
FT /\*tag= e  
FT /number= 2  
FT exon 5763..6019  
FT /\*tag= f  
FT /number= 3  
FT intron 6020..6104  
FT /\*tag= g  
FT /number= 3  
FT exon 6105..6238  
FT /\*tag= h  
FT /number= 4  
FT exon 6239..11091  
FT /\*tag= i  
FT /number= 4  
FT exon 11092..11247  
FT /\*tag= j  
FT /number= 5

WO2004077060-A2.  
XX  
XX 10-SEP-2004.  
XX  
XX 26-FEB-2004; 2004WO-CA000280.  
XX  
XX 27-FEB-2003; 2003US-0451382P.  
XX (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
XX Diamandis EP, Petraki CD;  
XX  
XX WPI; 2004-662077/64.

DR P-PSDB; ADR72621.  
XX  
XX Detecting kallikrein polypeptides/polynucleotides associated with renal  
FT cell carcinoma in a patient, for diagnosing/treating the disease,  
PT comprises detecting /identifying kallikrein polypeptides/polynucleotides  
XX in a sample.  
XX  
PS Example 1; SEQ ID NO 3; 53pp; English.  
XX  
CC The invention relates to a novel method for detecting kallikrein  
CC polypeptides, or the polynucleotides encoding them, associated with renal  
CC cell carcinoma. The method comprises obtaining a sample from a patient  
CC and detecting kallikrein polypeptides, or their encoding polynucleotides,  
CC where the kallikrein polypeptides are selected from the group consisting  
CC of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The  
CC detected amounts of the kallikrein polypeptides are compared to standard  
CC amounts. The molecules of the invention demonstrate cytostatic activity  
CC whilst the methods and kit may be useful for detecting, characterising,  
CC preventing and treating renal cell carcinoma. Furthermore, the methods  
CC may be useful for evaluating the probability of the presence of malignant  
CC or pre-malignant cells and for detecting and quantitating tumour growth  
CC and cancer metastasis. Finally, the methods may be utilised to confirm  
CC the absence or removal of all tumour tissue following surgery, cancer  
CC chemotherapy and/or radiation therapy and to monitor cancer chemotherapy  
CC and tumour reappearance. The current sequence is that of the human  
CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the  
CC invention which encodes a secreted serine protease and is located at  
CC chromosome 19q13.4.  
XX  
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;  
Query Match 100.0%; Score 5000; DB 13; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGGCCACAGTGAAGGCAAGAGAGAGAGTGTGAGAGTCCCTCTCAAGTGGCTTGAGTC 60  
Db 1 GGGCCACAGTGAAGGCAAGAGAGAGTGTGAGAGTCCCTCTCTCAAGTGGCTTGAGTC 60  
Qy 61 TCCCTGCTCTAAATGCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
Db 61 TCCCTGCTCTAAATGCAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
Qy 121 AAGAAAG 180  
Db 121 AAGAAAG 180  
Qy 181 ACAGAGAGCCTGGGACACACAGGAGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAG 240  
Db 181 ACAGAGAGCCTGGGACACACAGGAGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAG 240  
Qy 241 ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGAGATTAAACAGAGTCCACAGATACACGC 300  
Db 241 ACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGAGATTAAACAGAGTCCACAGATACACGC 300  
Qy 301 AAAGGGGACAGACAGAGTTTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTT 360  
Db 301 AAAGGGGACAGACAGAGTTTTCAGGGTGTGTCTATGATCATCTTCTTTTTTTTTTTT 360  
Qy 361 TTTTCTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCAGCGCTGAGTGGCGGG 420  
Db 361 TTTTCTTTTTTTTTTTTGGAGCGGAGTCTCGCTCTGTGCGCCAGCGCTGAGTGGCGGG 420  
Qy 421 GATCTCGGCTCACTGCAAGCTCCGCTCTCCGGGTTCACGCCATTCTCTGCTCAGGCTC 480  
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCTCCGGGTTCACGCCATTCTCTGCTCAGGCTC 480  
Qy 481 CCAAGTAGCTGGGACTACAGCGCGCCGCCACTACAGCGCGCGCTAATTTTTTTTATTTTA 540  
Db 481 CCAAGTAGCTGGGACTACAGCGCGCGCCGCCACTACAGCGCGCGCTAATTTTTTTTATTTTA 540  
Qy 541 GTAGAGAGCGGGTTTACCGCTTTTAGCGGGAGTGGCTCGATCTCTCGACTCGTGATCC 600  
Db 541 GTAGAGAGCGGGTTTACCGCTTTTAGCGGGAGTGGCTCGATCTCTCGACTCGTGATCC 600



Qy	2761	GTGTGCTGCTTGCAGGCGAATTATGTGATTTGGCTGTGAGTGTGACGTTATGGATGCCCCGTA	2820
Db	2761	GTGTGCTGCTTGCAGGCGAATTATGTGATTTGGCTGTGAGTGTGACGTTATGGATGCCCCGTA	2820
Qy	2821	TTTGTGACCGTGTCACTACCTTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG	2880
Db	2821	TTTGTGACCGTGTGACTTACCTTGAAGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTGTG	2880
Qy	2881	TCTGTGTGAGCCCGTGTAAATGCTACTGTATGTGTGATGTGGTGCAGCTGTGTGTCTGGAGT	2940
Db	2881	TCTGTGTGAGCCCGTGTAAATGCTACTGTATGTGTGATGTGGTGCAGCTGTGTGTCTGGAGT	2940
Qy	2941	TTCTGTCTCTGCTTGGAGGATAGAGGGTGCAGGGTAGCTATCTCTGGAGATGGGTGC	3000
Db	2941	TTCTGTCTCTGCTTGGAGGATAGAGGGTGCAGGGTAGCTATCTCTGGAGATGGGTGC	3000
Qy	3001	CAGGTGACTGACTTGTGCACTGTGCCCTGTGTGCAGAAAGATPATGTGSCAGTCTGAAATC	3060
Db	3001	CAGGTGACTGACTTGTGCACTGTGCCCTGTGTGCAGAAAGATPATGTGSCAGTCTGAAATC	3060
Qy	3061	TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC	3120
Db	3061	TGTGCACACACGGCATCTGTGCGTGGCACTGAGACACTGTGGATGAGGGTGTGCGATCCC	3120
Qy	3121	GCTAGGCTGCCCGGAGCGTGTGTAATGATGTGTGTCACAGAGCTGTATGTGTAGCTGCACCTGTG	3180
Db	3121	GCTAGGCTGCCCGGAGCGTGTGTAATGATGTGTGTCACAGAGCTGTATGTGTAGCTGCACCTGTG	3180
Qy	3181	GAGCAACATGGGCGTCTGTGCAGAACTGGTGTGTCACAGGCGCCATCAGCAAGGGTAA	3240
Db	3181	GAGCAACATGGGCGTCTGTGCAGAACTGGTGTGTCACAGGCGCCATCAGCAAGGGTAA	3240
Qy	3241	GCCTGTGTTCTTGGGCTGAGTTCGTGAATGATGTGTGTCACAGGCGCCATCAGCAAGGGTAA	3300
Db	3241	GCCTGTGTTCTTGGGCTGAGTTCGTGAATGATGTGTGTCACAGGCGCCATCAGCAAGGGTAA	3300
Qy	3301	GAACAGGCGGGCGCGTGGCTCAACAGGCGCCATCAGCAAGGGTAA	3360
Db	3301	GAACAGGCGGGCGCGTGGCTCAACAGGCGCCATCAGCAAGGGTAA	3360
Qy	3361	GGCGGATCACTGAGTGTGGGAGATCAGGCGCAGCTGACCAACATGGAGAACCCCGTCT	3420
Db	3361	GGCGGATCACTGAGTGTGGGAGATCAGGCGCAGCTGACCAACATGGAGAACCCCGTCT	3420
Qy	3421	CTACTAAAAATACAAAAATTAGCTGTGTGTGGCGCGTCCCTGTAAATCCACAGCTACTC	3480
Db	3421	CTACTAAAAATACAAAAATTAGCTGTGTGTGGCGCGTCCCTGTAAATCCACAGCTACTC	3480
Qy	3481	GGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGGAGGTGGAGTTCGGTGAAGCGGAG	3540
Db	3481	GGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGGAGGTGGAGTTCGGTGAAGCGGAG	3540
Qy	3541	TCGGCCATTGCACTCAGCCTGGGCAACAGAGCGAAACTCCGTCTCGAAAGAAAAA	3600
Db	3541	TCGGCCATTGCACTCAGCCTGGGCAACAGAGCGAAACTCCGTCTCGAAAGAAAAA	3600
Qy	3601	GAIAAAAAAGGCTAAGAACCGTGAATGGGCAACGGAGGACTGATGTGGAGTGGGG	3660
Db	3601	GAIAAAAAAGGCTAAGAACCGTGAATGGGCAACGGAGGACTGATGTGGAGTGGGG	3660
Qy	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGAGAGGAGGATTTGACAGGATTCAGAAAGGCAT	3720
Db	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGAGAGGAGGATTTGACAGGATTTGAGAGGAGCAT	3720
Qy	3721	GTTTTTCATCTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTTCCATGTGGCCCCCTAAG	3780
Db	3721	GTTTTTCATCTGAGAAATTCAGAAACCTAGGCTGTCTTCCCTTCCATGTGGCCCCCTAAG	3780
Qy	3781	CTGAGCCCTTCTTCTGCTTTCGGAACCTAGCTCCGCCCATGAGCTCTGAC	3840
Db	3781	CTGAGCCCTTCTTCTGCTTTCGGAACCTAGCTCCGCCCATGAGCTCTGAC	3840
Qy	3841	CCACCTCTTTCCTCAACACCGCCCTAGGCGAGACTCTAGTGGACCCGCCCTAAGGCCA	3900

Db	3841	 CCACCTCTTTCTCTCAACACGCCCCCTAGGCCAGACTCTAGTGGACCCCGCCCTTAAGGCCCA	3900
Qy	3901	CACCCCTTTTGGCGCAGGCTCCACCCCTATTCTGTGGGTACCTTTCTAGAACCCCTTCAA	3960
Db	3901	CACCCCTTTTGGCGCAGGCTCCACCCCTATTCTGTGGGTACCTTTCTAGAACCCCTTCAA	3960
Qy	3961	AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTCTCCAGGCTGGAG	4020
Db	3961	AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTCTCCAGGCTGGAG	4020
Qy	4021	TGCAGTGGCGTGATCTCGGCTCACTGGAACCTCTGCCTCCAGGTTCAAGTGATCTCGT	4080
Db	4021	TGCAGTGGCGTGATCTCGGCTCACTGGAACCTCTGCCTCCAGGTTCAAGTGATCTCGT	4080
Qy	4081	GCCTCCACCTCTCTGAGTAGCTGGGATTTACAGGTGCGCGCACACAGCCCTGGCTAAATTTT	4140
Db	4081	GCCTCCACCTCTCTGAGTAGCTGGGATTTACAGGTGCGCGCACACAGCCCTGGCTAAATTTT	4140
Qy	4141	GTGTCTTTTAGTAGAGACAGGTTTTTCACTTTTGGCCAGGCTGGTCTCAAATCTCCAAACC	4200
Db	4141	GTGTCTTTTAGTAGAGACAGGTTTTTCACTTTTGGCCAGGCTGGTCTCAAATCTCCAAACC	4200
Qy	4201	TCAGGTGATCCGGCCACCTCGGCTCCAGAGTGTGGGGTTACAGGCTGAGCCACCGC	4260
Db	4201	TCAGGTGATCCGGCCACCTCGGCTCCAGAGTGTGGGGTTACAGGCTGAGCCACCGC	4260
Qy	4261	CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTCTAACNTGTAACTCTGACCCCTGGCC	4320
Db	4261	CCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTCTAACNTGTAACTCTGACCCCTGGCC	4320
Qy	4321	CTAACTAAGTCAATTTCCAAACCCCTTCTCGCTCCAGCCCTGACCCCACTCACTGAGGCC	4380
Db	4321	CTAACTAAGTCAATTTCCAAACCCCTTCTCGCTCCAGCCCTGACCCCACTCACTGAGGCC	4380
Qy	4381	TGACCCCACTCTCTTGAGACCAAGTTCCATCCCTAAGGCCCTGTCCCTCCCATCCCCAG	4440
Db	4381	TGACCCCACTCTCTTGAGACCAAGTTCCATCCCTAAGGCCCTGTCCCTCCCATCCCCAG	4440
Qy	4441	GCTCCAGCCCCACAGCTTTGGCAGTACCCCTGAGCTTGCAGGGAATCCTGTACCCCAAT	4500
Db	4441	GCTCCAGCCCCACAGCTTTGGCAGTACCCCTGAGCTTGCAGGGAATCCTGTACCCCAAT	4500
Qy	4501	TTTACCCTCACATGTAGTTCTAGCCAAATTTCCAGGAATCTGTGAGGTCCAGTTAGAGTCCA	4560
Db	4501	TTTACCCTCACATGTAGTTCTAGCCAAATTTCCAGGAATCTGTGAGGTCCAGTTAGAGTCCA	4560
Qy	4561	GTAACCTCACTGAGCCTTGGGCTCTGCTTGAGCTTGAGCCTGGGCTTTGAGAGGTGCCA	4620
Db	4561	GTAACCTCACTGAGCCTTGGGCTCTGCTTGAGCTTGAGCCTGGGCTTTGAGAGGTGCCA	4620
Qy	4621	CTCTTATTTCTCAGGCCCTGCCCTGCCCTCAGCATGTCCAGACACCACTCTAGCT	4680
Db	4621	CTCTTATTTCTCAGGCCCTGCCCTGCCCTCAGCATGTCCAGACACCACTCTAGCT	4680
Qy	4681	GGTCTGGCTCTTTGAGTCTGAAACCCACCCAGCCCAAGCCCGCTCTGAGCCCGCCG	4740
Db	4681	GGTCTGGCTCTTTGAGTCTGAAACCCACCCAGCCCAAGCCCGCTCTGAGCCCGCCG	4740
Qy	4741	CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCAAACAATGATGTTTCTGTGACCAACC	4800
Db	4741	CAACCCATTTTCGTTTCCAGAGCATGTTCTCGCCAAACAATGATGTTTCTGTGACCAACC	4800
Qy	4801	CCTCTAACACCGTGCCCTCTGGAGACAAACAGAGCTGGGGCGGGGAGAGCG	4860
Db	4801	CCTCTAACACCGTGCCCTCTGGAGACAAACAGAGCTGGGGCGGGGAGAGCG	4860
Qy	4861	CCGGTCCGATGACAGCAGCCGATCATCAATGGATCCGATCCGATATGCCACACC	4920
Db	4861	CCGGTCCGATGACAGCAGCCGATCATCAATGGATCCGATCCGATATGCCACACC	4920
Qy	4921	AGCCGTGGCAGGCCGGCTTCTGCTAAGGCCCAACACAGCTCTACTCGCGGGCGGTGG	4980

Db 4921 AGCGTGGCGGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGG 4980  
Qy 4981 TGCATCCACAGTGGCTGCTC 5000  
Db 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 4  
ADR72875  
ID ADR72875 standard; DNA; 11570 BP.  
XX ADR72875;  
XX 02-DEC-2004 (first entry)  
XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.  
XX kallikrein 5; tumour marker; ovarian cancer;  
XX epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;  
KW KLK5; ds; gene.  
XX Homo sapiens.  
XX

Location/Qualifiers  
FH Key 2221..11247  
FT CDS /tag= b  
FT /product= "Human ovarian cancer-related tumour marker  
FT kallikrein 5 (hk5) protein"  
FT exon 2221..2293  
FT /tag= a  
FT /number= 1  
FT intron 2294..4761  
FT /tag= c  
FT /number= 1  
FT exon 4762..5023  
FT /tag= d  
FT /number= 2  
FT intron 5024..5762  
FT /tag= e  
FT /number= 2  
FT exon 5763..6019  
FT /tag= f  
FT /number= 3  
FT intron 6020..6104  
FT /tag= g  
FT /number= 3  
FT exon 6105..6238  
FT /tag= h  
FT /number= 4  
FT exon 6239..11091  
FT /tag= i  
FT /number= 4  
FT exon 11092..11247  
FT /tag= j  
FT /number= 5

W02004075713-A2.  
10-SEP-2004.  
26-FEB-2004; 2004WO-CA000281.  
26-FEB-2003; 2003US-0450406P.  
(MOUN ) MOUNT SINAI HOSPITAL.  
Diamandis EP;  
WPI: 2004-661815/64.  
P-PSDB; ADR72875.  
Kallikrein markers detection method for detecting ovarian cancer in  
patient, involves detecting kallikrein markers and CA125 in sample

obtained from patient, and comparing detected amounts with standard amounts.  
Example 2; SEQ ID NO 5; 102pp; English.  
The invention relates to a novel method for detecting a plurality of kallikrein markers associated with ovarian cancer. The method comprises obtaining a sample from a patient and detecting in the sample a plurality of kallikrein markers, and optionally carbohydrate antigen CA125, wherein the kallikrein markers are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11. The detected amounts of the kallikrein markers are compared with standard amounts. The method of the invention may be useful for detecting kallikrein markers associated with ovarian cancer in a patient and thus for detecting ovarian cancer, particularly epithelial ovarian carcinoma. The current sequence is that of the human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention which encodes a secreted serine protease and is located at chromosome 19q13.4.  
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;  
Query Match 100.0%; Score 5000; DB 13; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GGGCCACAGAGTGAAGGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAGTGGCTTGAGTC 60  
Db 1 GGGCCACAGAGTGAAGGCAAGAGAGAGTGTAGAGCTCCCTCTGCAAGTGGCTTGAGTC 60  
Qy 61 TCCCTTGCCTAAATGCAGGGAGAGGGAGGACAGAAAGACAGGGAAGAGGGGTGGG 120  
Db 61 TCCCTTGCCTAAATGCAGGGAGAGGGAGGACAGAAAGACAGGGAAGAGGGGTGGG 120  
Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGATACACAACTACAGAAACACAGAGAGAACAC 180  
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGATACACAACTACAGAAACACAGAGAGAACAC 180  
Qy 181 ACAGAGAGCTGGGACACAGGGACACAGAGTACAGAGAAAGAGAGATAGAGAAAG 240  
Db 181 ACAGAGAGCTGGGACACAGGGACACAGAGTACAGAGAAAGAGAGATAGAGAAAG 240  
Qy 241 ACAAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300  
Db 241 ACAAAATGGAGACACAGAGGTGTAAGAAAGAGAGATTAACAGAGTCCAGATACACGC 300  
Qy 301 AAAGGGCAGAGACACAGTTTTTCAGGGTGTGTATGATCATCTCTTTTTTTTTTTT 360  
Db 301 AAAGGGCAGAGACACAGTTTTTCAGGGTGTGTATGATCATCTCTTTTTTTTTTTT 360  
Qy 361 TTTTTTTTTTTTTTTTCAGAGCGGAGTCTCGCTCTGCGCCAGCTGGAGTGGCGG 420  
Db 361 TTTTTTTTTTTTTTTTCAGAGCGGAGTCTCGCTCTGCGCCAGCTGGAGTGGCGG 420  
Qy 421 GATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTACGCCATTCTCCTCCCTCAGCCTC 480  
Db 421 GATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTACGCCATTCTCCTCCCTCAGCCTC 480  
Qy 481 CCAAGTAGCTGGGACTACAGGCGCCCGCCACTAGCGCCGGCTAAATTTTTTGTATTTTA 540  
Db 481 CCAAGTAGCTGGGACTACAGGCGCCCGCCACTAGCGCCGGCTAAATTTTTTGTATTTTA 540  
Qy 541 GTAGAGACGGGGTTTACCGTTTTTACCGGGATTTAGCGGGATTTAGCGGGATTTTAT 600  
Db 541 GTAGAGACGGGGTTTACCGTTTTTACCGGGATTTAGCGGGATTTAGCGGGATTTTAT 600  
Qy 601 GCCCGCTCGGCTCCCAAGTGTCTGGATTACAGGCGTGGAGCCACCGCCCGGCCATG 660  
Db 601 GCCCGCTCGGCTCCCAAGTGTCTGGATTACAGGCGTGGAGCCACCGCCCGGCCATG 660  
Qy 661 ATCATCTTCTGACTATGCTGATGACAAAGTACCTAAAGCCATCAGACTCTACCCCTTA 720  
Db 661 ATCATCTTCTGACTATGCTGATGACAAAGTACCTAAAGCCATCAGACTCTACCCCTTA 720

Qy	721	AATATGAGT	TTGGGCCAGGCA	CGGTGGCTCATG	CTTGTAAT	TTCCAGCAT	TTTGGGAGGC	780
Db	721	AATATGAGT	TTGGGCCAGGCA	CGGTGGCTCATG	CTTGTAAT	TTCCAGCAT	TTTGGGAGGC	780
Qy	781	AGAGGTGGGT	GAAATCACTT	GAGGCCAGAGT	TTTGAGACCAGC	CTTGGCCAACTGGT	GCGTCAAA	840
Db	781	AGAGGTGGGT	GAAATCACTT	GAGGCCAGAGT	TTTGAGACCAGC	CTTGGCCAACTGGT	GCGTCAAA	840
Qy	841	CTCTGTCTT	TTACTTAAAAA	AAAAA	AAAAA	AAATCAGCCGGT	TCGTGGGGCAC	900
Db	841	CTCTGTCTT	TTACTTAAAAA	AAAAA	AAAAA	AAATCAGCCGGT	TCGTGGGGCAC	900
Qy	901	ACCTGTAAT	CCCCAGCTATG	CTGGAGGCTG	GAGGCACGAGT	TCATCTGAA	CCCTGGAGGC	960
Db	901	ACCTGTAAT	CCCCAGCTATG	CTGGAGGCTG	GAGGCACGAGT	TCATCTGAA	CCCTGGAGGC	960
Qy	961	GAGGTTG	CAGTGGGCCGAGAT	CAATCA	CCGCCCTCCAG	CTGGGCCACAGAG	CAAGACT	1020
Db	961	GAGGTTG	CAGTGGGCCGAGAT	CAATCA	CCGCCCTCCAG	CTGGGCCACAGAG	CAAGACT	1020
Qy	1021	CTGTCTCA	ATAATAATAA	AAAAA	AAAAA	AAATCAGCCGGT	TCGTGGGGCAC	1080
Db	1021	CTGTCTCA	ATAATAATAA	AAAAA	AAAAA	AAATCAGCCGGT	TCGTGGGGCAC	1080
Qy	1081	AAAAA	AAAAATGCTCA	AAAAATAG	ACGAGTGA	AAATAAGG	AAAAATAATATG	1140
Db	1081	AAAAA	AAAAATGCTCA	AAAAATAG	ACGAGTGA	AAATAAGG	AAAAATAATATG	1140
Qy	1141	AGAACTCT	AAAGTATATTTGA	CAAAATCA	TTTTCAGAA	CCCTTTAAAAA	AGAAAGAAATCA	1200
Db	1141	AGAACTCT	AAAGTATATTTGA	CAAAATCA	TTTTCAGAA	CCCTTTAAAAA	AGAAAGAAATCA	1200
Qy	1201	GGCATGA	AAAGACHGGGAG	GAACHGGGAG	CAGAAAC	CACTTGGTGGCC	AAAGGAGAA	1260
Db	1201	GGCATGA	AAAGACHGGGAG	GAACHGGGAG	CAGAAAC	CACTTGGTGGCC	AAAGGAGAA	1260
Qy	1261	ACAAGGCT	CTTAAGACACAG	GAGGAGGAG	GAGAGAGT	GTAGTGTAG	ACAGACAGACAG	1320
Db	1261	ACAAGGCT	CTTAAGACACAG	GAGGAGGAG	GAGAGAGT	GTAGTGTAG	ACAGACAGACAG	1320
Qy	1321	AAAAAGA	CAGAGAGAGAG	CACAGACAG	AGAGAGGCG	AGAGGGATAG	AAAGA	1380
Db	1321	AAAAAGA	CAGAGAGAGAG	CACAGACAG	AGAGAGGCG	AGAGGGATAG	AAAGA	1380
Qy	1381	GAGAGCGG	TGGAGAGAG	CACAGATAT	TGAGAGAG	CTCAGAA	AGATAGCCGAGGA	1440
Db	1381	GAGAGCGG	TGGAGAGAG	CACAGATAT	TGAGAGAG	CTCAGAA	AGATAGCCGAGGA	1440
Qy	1441	GAACCA	CAGAGAGAT	TGAAAGAACT	CTTGAGAAAAA	CCAGAGACA	AAAGATGAAAGAGG	1500
Db	1441	GAACCA	CAGAGAGAT	TGAAAGAACT	CTTGAGAAAAA	CCAGAGACA	AAAGATGAAAGAGG	1500
Qy	1501	AGTATC	GAGGATGAAC	CACAGTGGT	TGGAATAG	CAAAATG	CAGAGAAAGCAAGCA	1560
Db	1501	AGTATC	GAGGATGAAC	CACAGTGGT	TGGAATAG	CAAAATG	CAGAGAAAGCAAGCA	1560
Qy	1561	TCCAGG	CCCAAGATAG	TACCCAGAGT	TGGTGTAG	AAAGCCAGAT	CTTAAAGCTTGGGG	1620
Db	1561	TCCAGG	CCCAAGATAG	TACCCAGAGT	TGGTGTAG	AAAGCCAGAT	CTTAAAGCTTGGGG	1620
Qy	1621	AGGCAGG	GAAGGGCTG	CGCTGGCT	TCGGAGAC	CCCTTCCCA	TTCTCCGGGCCAGGGAG	1680
Db	1621	AGGCAGG	GAAGGGCTG	CGCTGGCT	TCGGAGAC	CCCTTCCCA	TTCTCCGGGCCAGGGAG	1680
Qy	1681	GTAGG	GAGTGCAT	TCCGGA	CTGGGTGGGGG	TGCTCTGGGGG	TGGAGATAGGGGGAGCA	1740
Db	1681	GTAGG	GAGTGCAT	TCCGGA	CTGGGTGGGGG	TGCTCTGGGGG	TGGAGATAGGGGGAGCA	1740
Qy	1741	GGAGG	AGCTAT	TGTAAG	CCCGATAG	GCACCTCAT	TGCCCCGGAATGTGCC	1800
Db	1741	GGAGG	AGCTAT	TGTAAG	CCCGATAG	GCACCTCAT	TGCCCCGGAATGTGCC	1800

QY	1801	CAGTGGGTGTTATTAACATCAGGCCCGGTGCCCCAGAGACCACAGHAGGAGGCAGTAGTGCCACAGA	1866
DB	1801	CAGTGGGTGTTATTAACATCAGGCCCGGTGCCCCAGAGACCACAGHAGGAGGCAGTAGTGCCACAGA	1866
QY	1861	AGGCACAGGCCTGAGAAGTCCTCGCGCTGAGCTCGGGAGCAAAATCCCCCACCCCCTACTCTG	1920
DB	1861	AGGCACAGGCCTGAGAAGTCCTCGCGCTGAGCTCGGGAGCAAAATCCCCCACCCCCTACTCTG	1920
QY	1921	GGGACAGGGCAAFTGAGACCTGTGTGAGGGTGGCTCAGCAGCGCAGGAGGAAGAGAGGTGTCT	1980
DB	1921	GGGACAGGGCAAFTGAGACCTGTGTGAGGGTGGCTCAGCAGCGCAGGAGGAAGAGAGGTGTCT	1980
QY	1981	GTCGGTCTCGACCCACATCTTTCTCTGTCCCCTCTGCCCCTGCTCGGANGCTGTACA	2040
DB	1981	GTCGGTCTCGACCCACATCTTTCTCTGTCCCCTCTGCCCCTGCTCGGANGCTGTACA	2040
QY	2041	CTCCATCTTTCTGAATTCTATAGTGTGCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGCTC	2100
DB	2041	CTCCATCTTTCTGAATTCTATAGTGTGCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGCTC	2100
QY	2101	CTTGTGGTTCCTCTCTA CTTGGGGAATAAGTGTAGGGAGGGAGGGAAGTGGGTTAAGG	2160
DB	2101	CTTGTGGTTCCTCTCTA CTTGGGGAATAAGTGTAGGGAGGGAGGGAAGTGGGTTAAGG	2160
QY	2161	GCTCCCCGATCGCTGGGCTCCCAACCTCTGAATTCGCCATCCAGGTGAGCGGCC	2220
DB	2161	GCTCCCCGATCGCTGGGCTCCCAACCTCTGAATTCGCCATCCAGGTGAGCGGCC	2220
QY	2221	ATGGCTACAGAACAGACCCCTCGATGTGGTGSCTCTGTGCTCTGATACAGCCTTGCCT	2280
DB	2221	ATGGCTACAGAACAGACCCCTCGATGTGGTGSCTCTGTGCTCTGATACAGCCTTGCCT	2280
QY	2281	CTGGGGTCTACAGTAACCAAGAACTCTGGGGTGGGAGGGTGTGGGATTTGGGAGGACTGT	2340
DB	2281	CTGGGGTCTACAGTAACCAAGAACTCTGGGGTGGGAGGGTGTGGGATTTGGGAGGACTGT	2340
QY	2341	CTCTGCGGCATAGAGGCCCTGTCCCTCGGGAACTGTGTGAGCCTTGCGCATGACTCCGG	2400
DB	2341	CTCTGCGGCATAGAGGCCCTGTCCCTCGGGAACTGTGTGAGCCTTGCGCATGACTCCGG	2400
QY	2401	GACCGGGTAATGTGAGTCTCTCTGTACTTGTGCTGTGGCATCGTATGTGGCCCTGT	2460
DB	2401	GACCGGGTAATGTGAGTCTCTCTGTACTTGTGCTGTGGCATCGTATGTGGCCCTGT	2460
QY	2461	GACTGCCACGGTGTGTCTCGGGAGGGGATGCTTTTTCCCATATCAGGTGACTGTGCCG	2520
DB	2461	GACTGCCACGGTGTGTCTCGGGAGGGGATGCTTTTTCCCATATCAGGTGACTGTGCCG	2520
QY	2521	CAGGTGCACTGACCCCTTTGAGCGTGTGTGTGGTGTGTGATGTGTGTCATTTAAG	2580
DB	2521	CAGGTGCACTGACCCCTTTGAGCGTGTGTGTGGTGTGTGATGTGTGTCATTTAAG	2580
QY	2581	ATTGTGTGGCTCCAAGCTGTGGGTGAATGCAATGACATGCGGGGTGTACTGT	2640
DB	2581	ATTGTGTGGCTCCAAGCTGTGGGTGAATGCAATGACATGCGGGGTGTACTGT	2640
QY	2641	GTGTTTGGCTGTGTGGTCACTTGGCATTTATATGACTGCAAGTATCTGCAAGTTCCCTG	2700
DB	2641	GTGTTTGGCTGTGTGGTCACTTGGCATTTATATGACTGCAAGTATCTGCAAGTTCCCTG	2700
QY	2701	TCCCTGAGGTCCCGGATTCGTCGCAAAAAGTGTTCATCCACATGGAAGCTGTGACT	2760
DB	2701	TCCCTGAGGTCCCGGATTCGTCGCAAAAAGTGTTCATCCACATGGAAGCTGTGACT	2760
QY	2761	GTGTGCTGCTTGACGGGATTTATGTGATTTGTGCTCAGTGTGACGTTATGATGCCCGTA	2820
DB	2761	GTGTGCTGCTTGACGGGATTTATGTGATTTGTGCTCAGTGTGACGTTATGATGCCCGTA	2820
QY	2821	TTTGTGACCGGTGACTPACCTGAGGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG	2880
DB	2821	TTTGTGACCGGTGACTPACCTGAGGCTCTGTGTAGGGGTGACTGTATGTGACTGTGTG	2880
QY	2881	TCTGTGTAGGCCCGGTAAAAATGCTACTGTATGTGTGATGGTGTGACGCTGTGTCTGGAGT	2940

Db 2881 TCTGTGAGCCGTGTAATGCTACTATGTGTGATGGTGCAGCTGTGTGTCTGGAGT 2940  
Qy 2941 TTCTGTCTCTCCCTGGAGGATAGAGGTGCAGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
Db 2941 TTCTGTCTCTCCCTGGAGGATAGAGGTGCAGGGTAGCTATCTCTGGGAGATGGGTGC 3000  
Qy 3001 CAGTGACTGACTGCACTGTGCTGCTGTGTGCAAGAGATGTGGCAGTCTGAACATC 3060  
Db 3001 CAGTGACTGACTGCACTGTGCTGCTGTGTGCAAGAGATGTGGCAGTCTGAACATC 3060  
Qy 3061 TGTGCACACAGGCATCTGTGCTGGCACTGTGACACACTGTGGATGAGGGTGTGGATCCC 3120  
Db 3061 TGTGCACACAGGCATCTGTGCTGGCACTGTGACACACTGTGGATGAGGGTGTGGATCCC 3120  
Qy 3121 GCTAGGCTGCCCGGAGCGTGTGTACCTGGAGACAGAGCTGTATTTAGCTGCACCTGTG 3180  
Db 3121 GCTAGGCTGCCCGGAGCGTGTGTACCTGGAGACAGAGCTGTATTTAGCTGCACCTGTG 3180  
Qy 3181 GAGCAACATGGCGTGTCTGCAGAACTGTGCTGCTGTGGCTGTACTGCTTTGTGC 3240  
Db 3181 GAGCAACATGGCGTGTCTGCAGAACTGTGCTGCTGTGGCTGTACTGCTTTGTGC 3240  
Qy 3241 GCGTGGTCTTGGGTGAGTTCTGTAATGATGTGTGTGGTGTACTGCTTTGTGC 3300  
Db 3241 GCGTGGTCTTGGGTGAGTTCTGTAATGATGTGTGTGGTGTACTGCTTTGTGC 3300  
Qy 3301 GAACAGCGCGGCGGTGTGTACCGCTGTATCCAGCCCTTTGGGAGGCGGAGGCA 3360  
Db 3301 GAACAGCGCGGCGGTGTGTACCGCTGTATCCAGCCCTTTGGGAGGCGGAGGCA 3360  
Qy 3361 GCGGATCACTGAGTCTGAGGTCAGAGCCAGCTGTGACCAATGGAGAACCCCGTCT 3420  
Db 3361 GCGGATCACTGAGTCTGAGGTCAGAGCCAGCTGTGACCAATGGAGAACCCCGTCT 3420  
Qy 3421 CTACTAAAAATACAAAAATTTAGTGTGTGTGGTGTGCTGTATCCAGCTACTC 3480  
Db 3421 CTACTAAAAATACAAAAATTTAGTGTGTGTGGTGTGCTGTATCCAGCTACTC 3480  
Qy 3481 GGGAGCTGGGCGAGAAATCGTTGAAACCGGAGGTGAGGTTCGGTGTAGCCGAGA 3540  
Db 3481 GGGAGCTGGGCGAGAAATCGTTGAAACCGGAGGTGAGGTTCGGTGTAGCCGAGA 3540  
Qy 3541 TCGGCCATTCGACTCCAGCTGGGCAACAGAGCCGAACTCGCTCTCGAAAGAAAAAA 3600  
Db 3541 TCGGCCATTCGACTCCAGCTGGGCAACAGAGCCGAACTCGCTCTCGAAAGAAAAAA 3600  
Qy 3601 GAAAAAAGGTAAGAACCACTGTAATGGGCAACGGAGGACTGATGTGGATGGGC 3660  
Db 3601 GAAAAAAGGTAAGAACCACTGTAATGGGCAACGGAGGACTGATGTGGATGGGC 3660  
Qy 3661 ATGCATCTAGTCTGTAGTGTGTGAGAGGAGGATGACAGATTGAGAGGCAT 3720  
Db 3661 ATGCATCTAGTCTGTAGTGTGTGAGAGGAGGATGACAGATTGAGAGGCAT 3720  
Qy 3721 GTTTTCACTGAGAAATTCAGAACTAGGCTGTCTTCCCTCCATGTGCCCCCTTAAG 3780  
Db 3721 GTTTTCACTGAGAAATTCAGAACTAGGCTGTCTTCCCTCCATGTGCCCCCTTAAG 3780  
Qy 3781 CTGAGCCCTTTCTTCTGTGCTGTTCGGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840  
Db 3781 CTGAGCCCTTTCTTCTGTGCTGTTCGGAAACCTTAGCTCCGCCCATGAGCTCTGACC 3840  
Qy 3841 CCACCTCTTCTCAACAGGCCCTTAGGCGAGACTCTAGTGAACCCGCCCTAAGGCA 3900  
Db 3841 CCACCTCTTCTCAACAGGCCCTTAGGCGAGACTCTAGTGAACCCGCCCTAAGGCA 3900  
Qy 3901 CACCCCTTTGGGCGAGGCTCAACCCCTATTTCTGTGGTACCTTTCTAGAACCCCTTCAA 3960  
Db 3901 CACCCCTTTGGGCGAGGCTCAACCCCTATTTCTGTGGTACCTTTCTAGAACCCCTTCAA 3960  
Qy 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTCTCTCTCCAGGCTGGAG 4020

Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTCTCTCTCCAGGCTGGAG 4020  
Qy 4021 TGCAGTGGGTGATCTCGGCTCACTCAACCTCTGCTCTCCAGGCTCAAGTGAATCTCGT 4080  
Db 4021 TGCAGTGGGTGATCTCGGCTCACTCAACCTCTGCTCTCCAGGCTCAAGTGAATCTCGT 4080  
Qy 4081 GCCTCCACCTCTCTGAGTAGCTGGGATTTACAGGTGCGCGCCACACCGCTGGCTAAATTTT 4140  
Db 4081 GCCTCCACCTCTCTGAGTAGCTGGGATTTACAGGTGCGCGCCACACCGCTGGCTAAATTTT 4140  
Qy 4141 GTGTCTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGGTCTCAAACTCCCAACC 4200  
Db 4141 GTGTCTTTAGTAGAGACAGGGTTTACCTTTGTTGGCCAGGCTGGTCTCAAACTCCCAACC 4200  
Qy 4201 TCAAGTGTATCCGCCACCTCGGCTCCAGAGTGTGGGTGTACAGGCGTGAAGCCCGC 4260  
Db 4201 TCAAGTGTATCCGCCACCTCGGCTCCAGAGTGTGGGTGTACAGGCGTGAAGCCCGC 4260  
Qy 4261 CCCAGCCCAAGTCAAGAGCTCTTTATAGGAGACTCTAAACATGTAACCTGACCTGGCC 4320  
Db 4261 CCCAGCCCAAGTCAAGAGCTCTTTATAGGAGACTCTAAACATGTAACCTGACCTGGCC 4320  
Qy 4321 CTAACTAAGTCAATTCCTGCTCTCCAGCCCTGAGGCTGACCTGAGGCT 4380  
Db 4321 CTAACTAAGTCAATTCCTGCTCTCCAGCCCTGAGGCTGACCTGAGGCT 4380  
Qy 4381 TGACCCCACTTTTGAGAGCAAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCAG 4440  
Db 4381 TGACCCCACTTTTGAGAGCAAGTTCCATCCCTAAAGCCCTGGTCTCCCTCCCAATCCCAG 4440  
Qy 4441 GCTCCAGCCCCACACAGCTTTGGCACTACCCCTGAGCTGTGTCAGGAATCTGTACCCAAT 4500  
Db 4441 GCTCCAGCCCCACACAGCTTTGGCACTACCCCTGAGCTGTGTCAGGAATCTGTACCCAAT 4500  
Qy 4501 TTTATCCCTCACATGTAGTTCTAGCCAAATTCAGGAATCTGTGAGGTCAGTTAGAGTCCA 4560  
Db 4501 TTTATCCCTCACATGTAGTTCTAGCCAAATTCAGGAATCTGTGAGGTCAGTTAGAGTCCA 4560  
Qy 4561 GTAAACCTTACCTGAGCTGGGCTCTGCTGTGAGCTTGAGCTGGGCTTGAGAGGTGCCA 4620  
Db 4561 GTAAACCTTACCTGAGCTGGGCTCTGCTGTGAGCTTGAGCTGGGCTTGAGAGGTGCCA 4620  
Qy 4621 CTCTTATTTCTCAGAGGCTGCCCCCTGCCCTCAGCATGTGTCAGACACCCACCTCTAGCT 4680  
Db 4621 CTCTTATTTCTCAGAGGCTGCCCCCTGCCCTCAGCATGTGTCAGACACCCACCTCTAGCT 4680  
Qy 4681 GGTCTGGCCTTTGAGTCTGAAACCCCAACCCAGCCCAAGCCCGCTCTGAGCCCGCC 4740  
Db 4681 GGTCTGGCCTTTGAGTCTGAAACCCCAACCCAGCCCAAGCCCGCTCTGAGCCCGCC 4740  
Qy 4741 CAACCCATTTTCCGTTCAGAGCATGTTCTCGCCAAACATGATGTTTCTGTGACCAAC 4800  
Db 4741 CAACCCATTTTCCGTTCAGAGCATGTTCTCGCCAAACATGATGTTTCTGTGACCAAC 4800  
Qy 4801 CCTCTAAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGAAGACG 4860  
Db 4801 CCTCTAAACCGTCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGAAGACG 4860  
Qy 4861 CCGGTGCGATGACAGCAGCGCGCATCAATGGATCCGACTCGGATATGCACACCC 4920  
Db 4861 CCGGTGCGATGACAGCAGCGCGCATCAATGGATCCGACTCGGATATGCACACCC 4920  
Qy 4921 AGCGGTGGAGCGCGGCTGTTGCTAAGGCCCAACAGCTCTACTCGGGGCGGTGTTGG 4980  
Db 4921 AGCGGTGGAGCGCGGCTGTTGCTAAGGCCCAACAGCTCTACTCGGGGCGGTGTTGG 4980  
Qy 4981 TGCATCCAGTGGCTGCTC 5000  
Db 4981 TGCATCCAGTGGCTGCTC 5000



ACH80345 standard; DNA; 525 BP.  
ACH80345;  
29-JUL-2004 (first entry)  
Human genome derived single exon probe #13540.  
Human; probe; ss; gene expression; single exon probe; microarray;  
alternative splicing event; genomic alteration.  
Homo sapiens.  
US2003194704-A1.  
16-OCT-2003.  
03-APR-2002; 2002US-00029386.  
03-APR-2002; 2002US-00029386.  
(PENW/) PENN S G.  
(RANK/) RANK D R.  
(HANZ/) HANZEL D K.  
Penn SG, Rank DR, Hanzel DK;  
WPI; 2004-119264/12.  
New human genome-derived single exon nucleic acid probes useful for human  
gene expression analysis, for identifying or characterizing alternative  
splicing events, for assessing genomic alterations or as tools for  
surveying tissues.  
Claim 15; SEQ ID NO 13540; 80pp; English.  
The invention relates to a nucleic acid probe for measuring human gene  
expression, comprising any of the 27,400 fully defined nucleotide  
sequences in the specification, or their complements or fragments, and  
encoding at least 8 amino acids of any of the 6888 amino acid sequences  
fully defined in the specification. The probe is a single exon probe that  
hybridises under high stringency conditions to a nucleic acid molecule  
expressed in human cells or tissues. Also included are a spatially-  
addressable set of single exon nucleic acid probes for measuring human  
gene expression (comprising a plurality of single exon nucleic acid  
probes cited above, where each of the plurality of probes is separately  
and addressably isolatable or amplifiable from the plurality), a single  
exon microarray for measuring human gene expression, a method of  
measuring human gene expression, a vector comprising the single exon  
probe cited above, an ORF-encoded peptide comprising at least 8  
contiguous amino acids of any of the above-mentioned amino acid  
sequences (optionally with conservative amino acid substitutions), an  
isolated antibody that binds specifically to a peptide cited above,  
methods of selling and/or licensing single exon probes or microarrays to  
a customer desiring to measure gene expression, a method of providing  
human gene expression data by subscription, and a computer-readable  
storage medium which contains a database having a plurality of records  
(each record including data on the expression of a single exon probe  
cited above). The probe, methods and apparatus are useful in gene  
expression analysis. The probes may be used as tools for surveying  
tissues to detect the presence of expressed messages that contain their  
specific exon, or in constructing genome-derived single exon microarrays.  
In addition, the probes are used in identifying and characterising  
alternative splicing events, in detecting and characterising gross  
alterations in the genomic locus that includes their exon, in assessing  
smaller genomic alterations, in priming the synthesis of nucleic acids,  
or in expressing the ORF-encoded peptide. The present sequence is a human  
single exon probe of the invention. Note: The sequence data for this  
patent did not form part of the printed specification, but was obtained  
in electronic format directly from USPTO at  
seqdata.uspto.gov/sequence.html?DocID=20030194704  
Sequence 525 BP; 99 A; 178 C; 145 G; 103 T; 0 U; 0 Other;

Query Match 9.4%; Score 472; DB 12; Length 525;  
Best Local Similarity 100.0%; Pred. No. 4.3e-76;  
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 4529 TCCAGGAATCTGTGAGGTCAGATTAGAGTCCAGTAACCCCTACCTGAGCCTGGGCTCTGTGTC 4588  
Db 1 TCCAGGAATCTGTGAGGTCAGATTAGAGTCCAGTAACCCCTACCTGAGCCTGGGCTCTGTGTC 60  
Qy 4589 CTTGAGCTTGAGCCTGGGCTTGAGAGTGCCACTCTTATTCTCCAGGCCCTGCCCTGCC 4648  
Db 61 CTTGAGCTTGAGCCTGGGCTTGAGAGTGCCACTCTTATTCTCCAGGCCCTGCCCTGCC 120  
Qy 4649 CCCTCAGCATGTGAGACACACCCCTCTAGCTGTCTGGGCTCTTGGAGTCTGAAACCCAC 4708  
Db 121 CCCTCAGCATGTGAGACACACCCCTCTAGCTGTCTGGGCTCTTGGAGTCTGAAACCCAC 180  
Qy 4709 CCCAGGCCCAAGCCCGCCCTCTGAGCCCGCCCAACCCATTTTCCGTTCCAGAGCATGT 4768  
Db 181 CCCAGGCCCAAGCCCGCCCTCTGAGCCCGCCCAACCCATTTTCCGTTCCAGAGCATGT 240  
Qy 4769 TCTCGCCCAACAATGATGTTTCTGTGTGACACACCCCTCTTAAACCGTGCCTCTGGAGCAA 4828  
Db 241 TCTCGCCCAACAATGATGTTTCTGTGTGACACACCCCTCTTAAACCGTGCCTCTGGAGCAA 300  
Qy 4829 CCAGGACCTGGAGCTGGGCGCGGGGAGAGCCCGGTGCGATGACAGCAGCAGCCGCAT 4888  
Db 301 CCAGGACCTGGAGCTGGGCGCGGGGAGAGCCCGGTGCGATGACAGCAGCAGCCGCAT 360  
Qy 4889 CATCAATGATCCGACTGCGATATGCACACCCAGCCGCTGGCAGCCGCGCTGTTGCTAAG 4948  
Db 361 CATCAATGATCCGACTGCGATATGCACACCCAGCCGCTGGCAGCCGCGCTGTTGCTAAG 420  
Qy 4949 GCCCAACCACTCTACTGCGGGCGGTGTTGGTGATCCACAGTGGTGTCTC 5000  
Db 421 GCCCAACCACTCTACTGCGGGCGGTGTTGGTGATCCACAGTGGTGTCTC 472

RESULT 6  
ACN44942  
ID ACN44942 standard; DNA; 41454 BP.  
XX ACN44942;  
XX ACN44942;  
XX 18-NOV-2004 (first entry)  
XX Human genomic sequence hCG21040.  
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.  
XX Homo sapiens.  
XX WO2003073826-A2.  
XX 12-SEP-2003.  
XX 28-FEB-2003; 2003WO-US006235.  
XX 01-MAR-2002; 2002US-00087192.  
XX (SAGR-) SAGRES DISCOVERY.  
XX Morris DW;  
XX WPI; 2003-328604/31.  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
PT comprises a nucleotide sequence.  
XX Claim 1; SEQ ID NO 1642; Opp; English.  
XX The present invention relates to novel DNA and protein sequences which  
CC are associated with carcinomas. The sequences are useful for: (i) for























GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 567.479 Seconds  
(without alignments)  
14417.061 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_1\_5000

Perfect score: 5000

Sequence: 1 gggccagagtggaagcaag.....tgatccacagtggtgctc 5000

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgm2\_6/ptodata/1/ina/5A COMB.seq:\*
  - 2: /cgm2\_6/ptodata/1/ina/5B COMB.seq:\*
  - 3: /cgm2\_6/ptodata/1/ina/6A COMB.seq:\*
  - 4: /cgm2\_6/ptodata/1/ina/6B COMB.seq:\*
  - 5: /cgm2\_6/ptodata/1/ina/PTCUS COMB.seq:\*
  - 6: /cgm2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2245	44.9	10818	4	US-09-949-016-13583
C 2	444	8.9	246240	2	US-08-724-394A-20
C 3	444	8.9	246240	2	US-08-724-394A-21
C 4	444	8.9	246240	2	US-08-724-394A-22
5	441	8.8	97376	4	US-09-949-016-16093
6	421.2	8.4	22973	4	US-09-949-016-13644
C 7	417.6	8.4	120727	4	US-09-949-016-15787
C 8	417.6	8.4	120727	4	US-09-949-016-15788
9	415.4	8.3	166698	4	US-09-949-016-16038
10	414.2	8.3	79858	4	US-09-949-016-16080
C 11	414	8.3	60304	4	US-09-949-016-12218
C 12	414	8.3	60305	4	US-09-949-016-15791
C 13	411.2	8.2	125188	4	US-09-949-016-11980
C 14	405.6	8.1	33379	4	US-09-949-016-13861
C 15	404.8	8.1	50217	4	US-09-949-016-16067
C 16	404	8.1	58593	4	US-09-949-016-12232
17	402.6	8.1	53442	4	US-09-949-016-11921
18	402.6	8.1	53453	4	US-09-949-016-13370
C 19	402.2	8.0	64518	4	US-09-949-016-17289
C 20	400.8	8.0	55114	4	US-09-949-016-16792
C 21	400.8	8.0	87617	4	US-09-949-016-16551
22	400.2	8.0	81701	4	US-09-949-016-14891
23	400.2	8.0	94855	4	US-09-949-016-12264
24	399.8	8.0	71879	4	US-09-949-016-17465
C 25	399.6	8.0	130563	4	US-09-949-016-12273
C 26	399.6	8.0	131379	4	US-09-949-016-16050
27	399.4	8.0	39920	4	US-09-949-016-16316

#### ALIGNMENTS

##### RESULT 1

US-09-949-016-13583

; Sequence 13583, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13583

; LENGTH: 10818

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-13583

Query Match 44.9%; Score 2245; DB 4; Length 10818;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2756	TGACTGTGTGCTGCTTCGACGCGATTATGTGATTTGGCTGAGTGTGACGTTATGGATGC	2815
Db	1	TGACTGTGTGCTGCTTCGACGCGATTATGTGATTTGGCTGAGTGTGACGTTATGGATGC	60
Qy	2816	CCGTAATTTGTACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGACTGT	2875
Db	61	CCGTAATTTGTACCGTGTGACTACCTGAAGCTCTGTGTAGGGGTGACTGTATGACTGT	120
Qy	2876	GTGTGTCTGTGTGAGGCGGTGTAATGCTACTGTATGTGTGATGCTGACGTTGTGTCT	2935
Db	121	GTGTGTCTGTGTGAGGCGGTGTAATGCTACTGTATGTGTGATGCTGACGTTGTGTCT	180
Qy	2936	GGAGTTTCTGTCTGCTGCTGAGGAGTAGAGGGGTGACGGGTAGCTATCTCTGGGAGATG	2995
Db	181	GGAGTTTCTGTCTGCTGCTGAGGAGTAGAGGGGTGACGGGTAGCTATCTCTGGGAGATG	240
Qy	2996	GGTGCCAGGTGACTGACTTGCAGTGTGTGCTGTGTGTCAGAGAGATATGTGGCAGTCTGA	3055
Db	241	GGTGCCAGGTGACTGACTTGCAGTGTGTGCTGTGTGTCAGAGAGATATGTGGCAGTCTGA	300
Qy	3056	ACATCTGTGCACACACCGGCATCTGTGCGTGGCACTGAGACACTGTGGATAGGGGTGTGCG	3115





[illegible]

## RESULT 4

US-08-724-394A-22/c  
Sequence 22, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Kronmal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolif, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and CREW LLP





Db 17544 TTTAGTAGACAGAGGTTTACCGTGTAGCAGGATGGTCTTGTGATCTCTCGACTCTCGTG 17603  
QY 597 ATCCGGCCGCTCGGCTCCCAAGTGTCTGGGATTACAGCGTGAGCCACCGCGCCGGC 656  
Db 17604 ATCCACTGCTCAGCTCCCAAGTGTCTGGGATTACAGCGTGAGCCACCGCTGAC 17663  
QY 657 CA-----TGATCATCTTCTTGACTATGCTGATGTA-----CAAGTACCTTAAAGCCA 703  
Db 17664 CAACTTTATTTCAGTCTTGGCTGGGCTTTTATTAACTTACTATAGAAATGTAACATA 17723  
QY 704 TCAGACTTACCTTTAAATATGCAAGTTGGGCGAGGACCGTGGCTCATGCTCTGTAAT 763  
Db 17724 CACAAAGTAGAATAGAGAATAGTATGATGGGCGAGTGCAGTGGCTCACACCTGTAATC 17783  
QY 764 CCAGCACTTTGGGAGGCGAGAGTGGTGAATCATCTTGAGCCAGGAGTTTGACACCAAGC 823  
Db 17784 CTAGCACTTTGGGAGGCTGAGGTGGGTGGATTCCTAGAGCTTAGGAGTTCTAGACGAGC 17843  
QY 824 TGCCCAACATGTGAAACTCTGCTTTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGC 883  
Db 17844 TGCCCAACATGTGAAACCCGCTCTACT-----AAAAATACAAAAATCGGC 17892  
QY 884 CGGGTGTGCGGGGCAACCTGTAATCCAGCTATGCTGGAGCTGAGGACGAGATCA 943  
Db 17893 TGGATGTGGTGGGCGACACCTGTAATCTCAGCTACTCGGAGGCTGAGGCGAGAATCA 17952  
QY 944 CTTGAACCTTGGAGGCGGAGGTTCAGTGGCGCGAGATCATCATCGGCCCTCAGCCCTG 1003  
Db 17953 CTTGAACCTTGGAGGCGGAGGTTCAGTGGCGCGAGATCATCGGCCCTGTTGTTCTAGCCTG 18012  
QY 1004 GCGCAGAGCAAGACTCTGCTCTCAATAATAAAT 1039  
Db 18013 GGCAACAGCGAGACTGCTCTCAAAAAAAAAAACT 18048

RESULT 7

US-09-949-016-15787/c  
; Sequence 15787, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15787  
; LENGTH: 120727  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(120727)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15787

Query Match 8.4%; Score 417.6; DB 4; Length 120727;  
Best Local Similarity 80.5%; Pred. No. 5.8e-87;  
Matches 546; Conservative 0; Mismatches 109; Indels 23; Gaps 4;  
QY 362 TTTTATTTTTTTTGGAGCGAGTCTGCTCTGCGCCAGGCTGAGTGCAGTGGCGGG 421  
Db 32483 TTCTTTTTTTTTTGGAGCGAGTCTGCTCTGCTCACCAGGCTGGAGTGCAGTGGCG 32424

QY 422 ATCTCGSCTCACTGCAAGCTCCGCTCCCGGGTTCA CGCAATCTCTCTGCTCAGCTCC 481  
Db 32423 ATCTCGSCTCACTGCAAGCTCCGCTCCCGGGTTCA CGCAATCTCTCTGCTCAGCTCC 32364  
QY 482 CAACTAGCTGGAGCTACAGCGCGCCCGCTAGCGCCGGCTAA -TTTTTTTGTATTTT 540  
Db 32363 CAACTAACAGGACTACAGCGCGCCCGCTAGCGCTGGCTAATTTTTTTTGTATTTT 32304  
QY 541 GTAGAGACGGGTTTTCACCGTTTATAGCGGGATGGCTCGATCTCTGACCTCTGATCC 600  
Db 32303 GTAGAGACGGGTTTTCACCTGTGTAGCAGGATGGCTTGTATCTCTGACCTCTGATCC 32244  
QY 601 GCCGCTCGGCTCCCAAAGTGTGGGATTTACAGCGGTGAGCCACCGCGCCGGCCATG 660  
Db 32243 ACCCACTCGGCTCCCAAAGTGTGGGATTTAGCGCGTGGAGCCACCGCGCCGGCCAGC 32184  
QY 661 ATCACTCTTGTACTATGCTGATGACAAAGTGTGAGCTTAAAGCATCAGACTCTACCTTTA 720  
Db 32183 CATGTCGGATTTCTGT-TGATGTTCTTCACTTACCTTGAAGGCCACGTGGAAAGAGGCTT-- 32127  
QY 721 AATATGCACTTTGGGCGAGCCAGCGTGGCTCATGCTGTAAATTCAGCACTTTGGGAGGC 780  
Db 32126 -----AGACTTGGCTGGTCAAGTGGCTCACACCTGTCTAGCACTTTGGGAGGC 32074  
QY 781 AGAGTGGGTGAATCACTTTAGCGCCAGGAGTTTGAAGCCAGCTTGGCAACATGCTGAAA 840  
Db 32073 CGAGGCGAGTGTACCTGAGCGCAGGAGTTTCAGCAAGTCTGACCAATGTTGAAA 32014  
QY 841 CTCTGTCTTACTATAAAAAAAAAAAAAAAAAAATCAGCGGGTGTCTGGGCGAC 900  
Db 32013 CCGCTCTCTACT-----AAAACTACAAAAATTAGCTGGCGGTGGTGGTGCAC 31966  
QY 901 ACTGTAAATCCAGCTATGCTGGAGCTGAGCGCAGGAGTCACTTGAACCTTGGAGCG 960  
Db 31965 ACTGTAAATCCAGCTACTCAGGAGCTGAGCGAGGAGTAATCTTTGAATCAGGAGGCA 31906  
QY 961 GAGGTTGAGTGGCGGAGATCACATCACCGCTCCAGCTGGGCGACAGAGCAAGACT 1020  
Db 31905 GAGGTTGAGTGGCGGAGTAATTTGGCCACTGCACTCCAGCTGGGTGACAGAGCCAGACT 31846  
QY 1021 CTGCTCAAAATAATAA 1038  
Db 31845 CTGCTCCAAAAACAAA 31828

RESULT 8

US-09-949-016-15788/c  
; Sequence 15788, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15788  
; LENGTH: 120727  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(120727)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15788

Query Match	8.4%;	Score 417.6;	DB 4;	Length 120727;
Best Local Similarity	80.5%;	Pred. No. 5.8e-87;		
Matches 546;	Conservative 0;	Mismatches 109;	Indels 23;	Gaps 4;
Qy	362	TTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTGTGCGCCAGAGCTGGAGTGCAGTGGCGGG	421	
Db	32483	TTCTTTTTTTTTTTTGGAGACGGAGTCTCGCTGTGTACCCAGAGCTGGAGTGCAGTGGTGG	32424	
Qy	422	ATTCGGCTCACTGCAAGCTCCGCTCCCGGGTTACGGCAATTCCTCTGCTTCAGCCCTCC	481	
Db	32423	ATTCGGCTCACTGCAAGCTCCGCTCCCGGGTTACGGCAATTCCTCTGCTTCAGCCCTCC	32364	
Qy	482	CAAGTAGCTGGGACTACAGCGCGCCGCGCACTACGCCCGGCTAA-TTTTTTTGTAATTTT	540	
Db	32363	CAAGTAACAGGGACTACAGCGCGCCGCGCACTACGCCCTGGCTAATTTTTTTGTAATTTT	32304	
Qy	541	GTAGACAGGGGTTTCACCGTTTTAGCGGGATGGCTCGATCTCCTGCACCTCGTGATCC	600	
Db	32303	GTAGACAGGGGTTTCACTGTGTTAGCCAGGATGGTCTTGATCTCCTGCACCTCGTGATCC	32244	
Qy	601	GCOCGGCTCGGCCTCCAAAGTCTGGGATTAACGGCGTGAGCCACCGCGCCGCGCAATG	660	
Db	32243	ACCCACCTCGGCCTCCCAAGTCTGGGATTAACGGCGTGAGCCACCGCGCCGCGCAATG	32184	
Qy	661	ATCATCTTTGTGACTATGCTGATGTGACAAAGTACTTAAAGCCATCAGACTCTACCCCTTA	720	
Db	32183	CATGTCGGATTCCTGT-TGATGTTTCTTCACTACTCTGAAGCCACCGTGAAGAGGCTT--	32127	
Qy	721	AATATGCAGTTTGGGCGAGCCACCGTGGCTCATGCTGTAAATTCAGACACTTTGGGAGGC	780	
Db	32126	-----AGCTTGGCTTGGTTCACGGTGGCTCACACCTGTCTCTAGACACTTTGGGAGGC	32074	
Qy	781	AGAGGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCCAGGCTTGCCCAACTGTTGAAA	840	
Db	32073	CGAGGCAGGTAGATCACTCGAGGCCAGGAGTTTCGAGACAAGTCTGCACCAACTGTGAAA	32014	
Qy	841	CTCTGTCTTTACTTAAAAAATAAAAAAATAAAAAAATCAGCCGGGTGTCTGTGGGGCAC	900	
Db	32013	CCCCGTCTCTACT-----AAAACTCAAAAAATTAGCTGGCGGTGGTGTGTGCAC	31966	
Qy	901	ACCTGTAATCCAGCTATGCTGGAGGCTCAGGACACGAGAGTCACTTTGAACCTCTGAGGCG	960	
Db	31965	ACCTGTAATCCAGCTACTCTAGGAGGCTGAGGAGAGGAATCTCTTGAACCTCAGGAGCA	31906	
Qy	961	GAGGTTGCAGTGGGCGGAGATCACATCACCGCCCTCCAGCTGGGCGCAGAGCAAGACT	1020	
Db	31905	GAGGTTGCAGTGAGCCAAAGATTGGCGCACTGCACTCCAGCCTGGGTGACAGAGCCAGACT	31846	
Qy	1021	CTGTCTCAATAAATAATAA	1038	
Db	31845	CTGTCTCCAAAAACAAAA	31828	

RESULT 9  
US-09-949-016-16038  
; Sequence 16038, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C0001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0

```
; SEQ ID NO 16038
; LENGTH: 166698
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(166698)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16038

Query Match      8.3%; Score 415.4; DB 4; Length 166698;
Best Local Similarity 76.2%; Pred. No. 2.2e-86;
Matches 560; Conservative 0; Mismatches 151; Indels 24; Gaps 3;

QY 328 TGGTGTCTATGATCATCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGAGTCT 387
Db 106710 TAGCTTGTTTTATTATTATTTATTTAATTTATTTATTTATTTATTTTGTAGATGAGTCT 106769

QY 388 CGCTCTCGCCCGCAGGCTGGAGTCGAGTCAGTGGCGGATCTCGGCTCACTGCAGACTCCGCT 447
Db 106770 TGCTCTGTGCGCCAGGCTGGAGTCGAGTCGAGTCATCTCTCTCACGCAAGCTCCGCT 106829

QY 448 CCGGGTTTCAGCCATTCTCTGCTCAGCTCCAAAGTAGCTGGAGCTACAGGCCGCCG 507
Db 106830 CCTGGTTTCAGCCATTCTCTGCTCAGCTCTCTGATTAGCTGGATTACAGGCCGCCG 106889

QY 508 CCATACGCCCGGCTAAATTTTTTTTGTATTTTTAGTAGAGACGGGTTTCACCGTTTTAGC 567
Db 106890 CCACCAGCCAGATAATTTTTTTGTATCTTTAGTAGAGTGGGTTTCACTGTGTAGC 106949

QY 568 CGGATGGCTCGATCTCTGACCTGTGATCGCGCGCTCGGCTCCCAAAGTGCTGG 627
Db 106950 CAGATGGTCTCAATCTCTGACCTGTGATCGCGCTCTCTCGGCTCCCAAAGTGCTGT 107009

QY 628 GATTACAGCGGTGAGCCACCGCGCCCGCCATGATCATCTTCTTCATCTATGCTGATGA 687
Db 107010 GATTACAGCGGTGAGCCACCGCGCCCGCTTGCTTTAAAACAACAACTATTACTG 107069

QY 688 CAAGTACTTAAGCC-----ATCAGACTCTACCTTTTAAATATGCAAGTTGGGC 736
Db 107070 AAAAAAAAAAATGTCCAATGAGATTGTGAAAGATTAGCATTAATAAATTAATAAGGC 107129

QY 737 CAGGCA -CCGTGGCTCATGCTGTAATTCAGACATTTGGAGCAGCAGGTTGGGTGATC 795
Db 107130 TTAAAGTGGTGGCTCAGCTGTAATCCAGCAATTTTAGGGCTGAACCGGGCAGATC 107189

QY 796 ACTTGAGGCAGGAGTTTGAGCCAGCCTGGCCAACATGGTGAAACTCTGCTTTACTTAA 855
Db 107190 ATTTGAGTCAGGAGTTTGAAACAGCCTGGCCAACTGTTGAACCCCGTCTCTACT-- 107247

QY 856 AAAAAAAAAAAAAAAAANAATCAGCGGGTGTGCTGGGGCACA CTCTGTAATCCCAGC 915
Db 107248 -----AAAAATCAAAAAATTAGCAGCGGTGGTGGCACA CAGCTGTAATCCCAGC 107297

QY 916 TATCTGGAGGCTCAGGCA CAGAGTCACTTGAACCTCGAGGCGGAGGTTGCA GTGGC 975
Db 107298 AACTCAGAGGCTCAGG CAGAGAATCGCTTGAACCCAGAAGGTGGAGCTTGCAGTGAGC 107357

QY 976 CGAGATCACATCA CCGCCCTCCAGCTGGGCGCAGACGAAGACTCTCTCTCAATAAAT 1035
Db 107358 TGAGATCATGCCATTGCACTCCAGCCTCCAGCCTGGGTGACAGACGAAGCTCCATCTCAAAAAA 107417

QY 1036 AAATAACAACGAA 1050
Db 107418 AAAAAAAAAAAAAAA 107432
```

RESULT 10  
US-09-949-016-16080  
; Sequence 16080, Application US/099459016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.









**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 2001.63 Seconds

(without alignments)

15667.668 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_1\_5000

Perfect score: 5000

Sequence: 1 gggccagagtgagggaag.....tgatccacagtggtgctc 5000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched:

6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	472	9.4	525	US-10-029-386-13540
2	440.2	8.8	41454	Sequence 13540, A
3	419	8.4	91352	Sequence 1642, Ap
4	418.4	8.4	115935	Sequence 4, Appli
5	417.6	8.4	26371	Sequence 241, App
6	417.6	8.4	91760	Sequence 1450, Ap
7	417.6	8.4	136726	Sequence 844, App
C				Sequence 244, App

8	416.2	8.3	379652	21	US-10-481-613-71	Sequence 71, Appl
9	415.2	8.3	17397	9	US-09-764-869-1945	Sequence 1945, Ap
10	415.2	8.3	17397	14	US-10-091-504-1945	Sequence 1945, Ap
11	415.2	8.3	17397	17	US-10-227-577-1945	Sequence 1943, Ap
12	415.2	8.3	19334	9	US-09-764-869-1943	Sequence 1943, Ap
13	415.2	8.3	19334	14	US-10-091-504-1943	Sequence 1943, Ap
14	415.2	8.3	19334	17	US-10-227-577-1943	Sequence 1944, Ap
15	415.2	8.3	19345	9	US-09-764-869-1944	Sequence 1944, Ap
16	415.2	8.3	19345	14	US-10-091-504-1944	Sequence 1944, Ap
17	415.2	8.3	19345	17	US-10-227-577-1944	Sequence 1944, Ap
C	415.2	8.3	29001	19	US-10-317-270-11	Sequence 11, Appl
C	415.2	8.3	12970	10	US-09-764-891-7689	Sequence 7689, Ap
20	414.6	8.3	111084	20	US-10-723-860-1627	Sequence 11267, Ap
21	414.2	8.3	212231	13	US-10-087-192-1126	Sequence 11267, Ap
C	412.8	8.3	3294	13	US-10-027-632-115723	Sequence 115723, Ap
C	412.8	8.3	3294	13	US-10-027-632-115724	Sequence 115724, Ap
C	412.8	8.3	3294	17	US-10-027-632-115724	Sequence 115724, Ap
C	412.8	8.3	3294	17	US-10-027-632-115724	Sequence 115724, Ap
26	411.2	8.2	13170	19	US-10-741-601-5744	Sequence 5744, Ap
27	411.2	8.2	13170	21	US-10-741-600-17903	Sequence 17903, A
28	410.4	8.2	10445	10	US-09-764-891-6380	Sequence 6380, Ap
29	409.6	8.2	13328	21	US-10-741-600-17627	Sequence 17627, A
C	409.6	8.2	17303	21	US-10-741-600-17720	Sequence 17720, A
C	409.6	8.2	44801	21	US-10-741-600-17966	Sequence 17966, A
C	409.6	8.2	50000	17	US-10-364-505-6	Sequence 6, Appli
C	409.6	8.2	50000	19	US-10-681-199-6	Sequence 6, Appli
C	407.6	8.2	187844	20	US-10-719-993-6883	Sequence 6883, Ap
C	407.6	8.2	196686	13	US-10-087-192-484	Sequence 484, App
C	406.2	8.1	12970	10	US-09-764-891-7689	Sequence 7689, Ap
C	405.8	8.1	4433	14	US-09-764-891-6830	Sequence 6830, Ap
C	405.8	8.1	4433	14	US-10-091-572-668	Sequence 668, App
C	405.8	8.1	126001	17	US-10-175-492-13	Sequence 13, Appli
40	405	8.1	32767	16	US-10-004-113-4	Sequence 4, Appli
41	405	8.1	32767	17	US-10-394-948-4	Sequence 4, Appli
C	405	8.1	67088	19	US-10-741-601-5704	Sequence 5704, Ap
C	405	8.1	67088	21	US-10-741-600-17804	Sequence 17804, A
C	403.8	8.1	104062	13	US-10-087-192-916	Sequence 916, App
C	403.6	8.1	1503841	9	US-09-795-668-1	Sequence 1, Appli

#### ALIGNMENTS

RESULT 1  
US-10-029-386-13540  
; Sequence 13540, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Hank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEOMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 13540  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR19.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.83  
; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUE 2.00e-34  
; OTHER INFORMATION: NT HIT: AFI35028.1, EVALUE 0.00e+00  
; OTHER INFORMATION: EST\_HUMAN HIT: BE388198.1, EVALUE 0.00e+00  
US-10-029-386-13540

Query Match 9.4%; Score 472; DB 16; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.5e-102;



















**THIS PAGE BLANK (uspto)**

Result No.	Query No.	Score	Match		Length	DB	ID	Description
C	1	397.6	8.0		2971	8	AF101960	AF101960 AF101960
	2	373.8	7.5		2971	8	AF101960	AF101960 AF101960
C	3	365	7.3		1715	3	CR592231	CR592231 full-length
C	4	362.6	7.3		2429	3	HSN805366	AL834319 Homo sapi
C	5	337.4	6.7		652	1	AU120416	AU120416 AU120416
C	6	337.2	6.7		1641	3	BC033224	BC033224 Homo sapi
C	7	336	6.7		2772	3	BC038630	BC038630 Homo sapi
C	8	334.4	6.7		698	5	BM999040	BM999040 UI-H-D10-
	9	333.8	6.6		528	5	AB012089	AB012089 Homo sapi
C	10	329	6.6		381	8	CC061638	CC061638 MUGQ CH25
	11	325.2	6.5		652	1	AU120416	AU120416 AU120416
C	12	324	6.5		701	2	BE744242	BE744242 601576428
C	13	324	6.5		5511	3	HSN807323	EX647179 Homo sapi
	14	322.8	6.5		6934	3	HSN807092	EX640943 Homo sapi
C	15	321.6	6.4		2772	3	BC038630	BC038630 Homo sapi
	16	318.6	6.4		5528	9	AB012089	AB012089 Homo sapi
C	17	315.6	6.3		709	9	AG010128	AG010128 Homo sapi
	18	315	6.3		710	6	CA428305	UI-H-DFO-
C	19	312.6	6.3		542	1	AI679782	AI679782 tu76f10. x
	20	312.2	6.2		495	1	AI963720	AI963720 wr65b04. x
C	21	311.8	6.2		2429	3	HSN805366	AL834319 Homo sapi
	22	311.4	6.2		3491	3	BC039100	BC039100 Homo sapi
C	23	309.6	6.2		1910	3	BC035771	BC035771 Homo sapi
	24	309.4	6.2		3491	3	BC039100	BC039100 Homo sapi

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84



```

ACCESSION      CR592231
VERSION        CR592231.1  GI:50473038
KEYWORDS       HTC; CNSLT_CDNA.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1715)
AUTHORS        Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE          Full-length cDNA libraries and normalization
JOURNAL        Unpublished
REMARK         Contact : Feng Liang Email : fliang@lifetech.com URL :
               http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
               Faraday Avenue
REFERENCE      2 (bases 1 to 1715)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
               BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)
               - Web : www.genoscope.cns.fr)
COMMENT        1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
               end enriched, double-strand cDNA was digested with Not I and cloned
               into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
               was normalized. Library was constructed by Life Technologies, a
               division of Invitrogen.
FEATURES       .
               . Location/Qualifiers
               .
               1..1715
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0D1023YE11"
                /tissue_types="Placenta Cot 25-normalized"
                /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      7.3%; Score 365; DB 3; Length 1715;
Best Local Similarity 71.5%; Pred. No. 5.1e-23;
Matches 551; Conservative 0; Mismatches 135; Indels 85; Gaps 2;

QY      355  TTTT TTTT TTTT TTTT TTTT TTTT TGAG CGAG TCTG CGTCTG TCGGCC CAGG CTGG AGTGCAG 414
Db       701  TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT
QY      415  TGG CGGG ATCT CGG CTCA CTG CAAG CTCC CGC TCTCC CGG GTTC AGC CATTTCT CTCTGC CTC 474
Db       761  TGG CGG ATCT CGG CTCA CTG GAG CTCTG TCTCT GGG TTTATG CCA TTTCTCTG CCTC 820
QY      475  AGCT TCCA AGT AGCT GGA CTAC AGC GCG CCG CCA CTAC GCG CGG CTA ATTTT TTTG TA 534
Db       821  AGCT TCCA AGT AGCT GGA CTAC AGC GCA CCC ACC CAC CAG CGC CGC CTA ATTTT TTTG TA 890
QY      535  TTTT TTAG TACAG CGGG GTTTC CCG TTTT TAG CGG GATG CCGT CGAT CTTCTG CACTCG 594
Db       881  TTTT TTAG TACAG ATG GGG GTTTC CCG TTTT TAG CCG GTTAG CCA TGA TGGTCTCC ATCTCT 940
QY      595  TGAT CGCG CGC CCG CTCTCG GCGTCC CAA AGTGTG GGA TTAC AGG GTG TAC GCG CAC CGC GCG 654
Db       941  TGAC TCGCTG CCG CTG GCG CCG CCG CCA AAG GGTG GGA TTAC AGG GTG TAC GCG CAC CGC 1000
QY      655  GCCA-----
Db       1001 GCCAC TTTCTCTTG ATA AAG ACCACTG ACTATG GGTG GTTCTG TGG CCA GTTTATAG 1060
QY      659  -----TGATCA TCTTTT GAC TATGCTG ATGTGAC AAGTAC CTAAGCC 702
Db       1061 GCTAAG CAC CCGTGTTC CTTCTG TCTGCTG CTAAG CAGG CATTG TGA CTTATAG GGTCTAAT 1120
QY      703  ATCAG ACTCTAC CCGT TTTAATATG CAG TTTTGG CCGAG CAC CGTGG CTTCA TCGCTG TGAAT 762
Db       1121 TTATG CATTAA GTTTAAG TCTCC ACTTAG GCGTGG CCGTGG TGGCTCAG GCTGTAAT 1180
QY      763  TCCAG CACTTTGG GAGCAG AGGTGG GTGAA TCA CTTGAG CCGCAG GAGTTTGAG ACCAGC 822

```





939	TTTTGTATTTTCAGTAGAGACAGGGTTTCACCTTGTGTAGCCAGGATGGTCTCGATCTCCT	880
588	GACCTGCTGATCCGCGCGCTCGGCTCCCAAAGTGTGGGATTTACAGGGGTGAGCCACC	647
879	GACCTCATGATCCACCTGCTCGGCTCCCAAAGTGTGGGATTTACAGGCATGAGCCACC	820
648	GGCGCGGGCCATCATCTCTTTGACTATGCTGATGTGACAAAGTACCTTAAAGCCATCAG	707
819	ATGCCAGGCCAACTCTCATCTTAATTGGACAATTTTAAGAAAGTATTTTACTCCAGAG	761
708	ACTTACCTCTTTAAATATGACGTTTGGGCGAGCCAGCGTGGCTCATGCTGTAAATTCAG	767
760	-----CTTATAAAACAAGCCTACAGCCAGGACAGTGGTGACGATGTAAATCCAG	709
768	CACCTTTGGGAGGAGAGGTGGGTGAATCACTTCAGGCCAGAGGTTTGAGACCAAGCCTGGC	827
708	CACCTTTGGGAGGCGAGGT--GGTGGATCACTTCAGGCCAGAGTTCAGACCAAGCCTGGG	650
828	CACATGTTGAAACTCTGCTCTTTACTTAAAAAATAAAAAAATAAAAAAATCAGCGCGG	887
649	CAACAGAGCGAGACTCAAACTCTACAGAAATTTAAAAAATTA-----GCCGG	602
888	TGTCGTGGGCGACACCTGTAAATCCAGCTATGCTGAGGCTGAGGCGACGAGATCACTTG	947
601	ACGTGGGGCACCCGCGCTGTAGTTCCAGTACTCAGGAG-----ATTGCTTG	555
948	AACCTTGGAGGGAGGTTTCAGTGGGCCAGAGTACATCACCGCCCTCCAGGCTGGCG	1007
554	AACCTGGAGTTCAAGGTTGACGTTGACGCTGACCATGCCACTCCACTCCAGGCTGGGTG	495
1008	ACAG 1011	
494	ACAG 491	

RESULT 7

BC038630/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC038630

Homo sapiens, Similar to hypothetical protein FLJ20489, clone

IMAGR:5263792, mRNA.

BC038630

BC038630.1

GI:24116283

HTC.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2772)

Strausberg, R.

Direct Submission

Submitted (15-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadano@systemsbiology.org](mailto:amadano@systemsbiology.org)

Anup Madan, Jessica Fahey, Erin Helton, Mark Ketterman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 73 Row: 1 Column: 4

This clone has the following problem: retained intron.

Location/Qualifiers

1..2772

source



```
Db      241  CTCCTGACCTCGTGATCTGCCCGCTCCACCTCCCAAGTGTGGATTACAGGGTGAG 300
Qy      643  CCACCGCGCGCCGATGATCATCTCTTGATCTATCTGCTGATGTGACAAAGTACCTAAAGCC 702
Db      301  CCACCATGCTCGGCCATTTATATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 360
Qy      703  ATCAGACTCTACCTTTAAATATGAGTTTGGGCCAGGCACCGTGGCTATGCTGTAAT 762
Db      361  GAAGACTGACAGACTACT-----GGTTGGCGGGCGGTGACTATATCTGAAT 414
Qy      763  TCCAGACTTTTGGAGGAGAGGTGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGC 822
Db      415  CCAGTACTTCAGGAAGCAAGAGGGTGAATCTTTGAGCCAGAAATTCAGACCAGG 474
Qy      823  CTGGCCAACTGGTGAATCTGTCTTACTTAAATTAATTAATTAATTAATTAATTAATTAAT 882
Db      475  CTGGCCAACTGGTGAATCTGTCTTACTTAAATTAATTAATTAATTAATTAATTAATTAAT 513
Qy      883  CCGGTGTGTCGGGCGACACCTGTAATCCAGCTATGCTGGAGGCTGAGGCACGAGTTC 942
Db      514  ---GGTGTGTCGGGCGACACCTGTCAGCTACTCGGGAGGCTGAGGTAATAGGCTC 570
Qy      943  ATTGAACCTCGAGGCGAGGTTCAGTGGGCCGAGATCAATCAACCGCCCTCCAGCT 1002
Db      571  ACTAGAACCTGGAGGTTAAGCTCGAGTGAGCCGAGATCTCGCCACTGTACTCCAGCT 630
Qy      1003  GGGCGACAGAGCA-AGACTCTCTCAATAATAATAATAATAATAATAATAATAATAATAA 1061
Db      631  GAGTGGCAGAGCANACCTATCTCANAAATAATTTTANNAAAGAAAAAATAATAATAA 690
Qy      1062  TTGTACC 1068
Db      691  TCGTCCC 697
```

## RESULT 9

AB012089 Homo sapiens gene, genomic survey sequence. 5528 bp DNA linear GSS 21-MAY-1998

LOCUS AB012089

DEFINITION Homo sapiens gene, genomic survey sequence.

ACCESSION AB012089

VERSION AB012089.1 GI:3149956

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Kurahashi,H., Sakamoto,M., Ono,J., Honda,A., Okada,S. and

Nakamura,X.

Molecular cloning of the chromosomal breakpoint in the Lif1 gene of

a patient with isolated lissencephaly and balanced t(8:17)

Hum. Genet. (1998) In press

2 (bases 1 to 5528)

Kurahashi,H.

Direct Submission

Submitted (09-MAR-1998) Hiroki Kurahashi, Biomedical Research

Center, Osaka University Medical School, Division of Clinical

Genetics, Department of Medical Genetics; 2-2 Yamadaoka, Suita,

Osaka 565-0871, Japan (E-mail:kurah@clgene.med.osaka-u.ac.jp,

Tel:81-6-879-3381, Fax:81-6-879-3389)

Location/Qualifiers

1..5528

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="17"

/map="17p13.3"

FEATURES

source

Query Match 6.6%; Score 331.8; DB 9; Length 5528;

Best Local Similarity 62.0%; Pred. No. 1.3e-19;

Matches 629; Conservative 0; Mismatches 347; Indels 39; Gaps 5;

ORIGIN

CC061638

LOCUS

381 bp DNA linear GSS 28-APR-2003

RESULT 10

CC061638/c

LOCUS

DEFINITION	MUGO_CH252P004Q3T7_M05_CD318_034 CHORI-252 Vervet Monkey Library Cercopithecus aethiops genomic clone CH252-4B9, genomic survey sequence.									
ACCESSION	CC061638									
VERSION	CC061638.1 GI:29790901									
KEYWORDS	GSS.									
SOURCE	Cercopithecus aethiops (African green monkey)									
ORGANISM	Cercopithecus aethiops									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Cercopithecinae; Cercopithecus.									
REFERENCE	1 (bases 1 to 381)									
AUTHORS	Minhas,R., Zhang,X., Dore,C., Villeneuve,A., Lepage,P., Forgetta,V., McKee,K., Ophoff,R.A., Fairbanks,J.A., Freimer,N.B., Ervin,F.R., Palmour,R.M., Hudson,T.J. and Dewar, K.									
TITLE	UCLA/MUGO/St-Kitts Vervet Monkey Mapping Project									
JOURNAL	Unpublished (2003)									
COMMENT	Contact: Dewar K McGill University and Genome Quebec Innovation Centre 740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1A4 Tel: 514 398 3311 x00089 Fax: 514 398 1795 Email: ken.dewar@mcgill.ca Plate: 4 row: B column: 9 Seq primer: T7 : TAATACGACTCATATAGGG Class: BAC ends.									
FEATURES	Location/Qualifiers									
source	1..381									
	/organism="Cercopithecus aethiops"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:9534"									
	/clone="CH252-4B9"									
	/sex="male"									
	/cell_type="white blood cell"									
	/dev_stage="Adult"									
	/clone_lib="CHORI-252 Vervet Monkey Library"									
	/note=Vector: pTRABAC2.1; Site 1: EcoRI; Site 2: EcoRI; Constructed by Michael Nefedov in Pieter de Jong's laboratory at BACPAC Resources, Children's Hospital in Oakland Research Institute."									
ORIGIN										
Query Match	6.6%; Score 329; DB 8; Length 381;									
Best Local Similarity	93.2%; Pred. No. 1.3e-18;									
Matches 355;	Conservative 0; Mismatches 25; Indels 1; Gaps 1;									
Qy	1649	GGAGACCCCTCCCATTTCTCGGCGCCAGGGAGTAGGAGTGCACATTCCCGACTCGGTGG	1708							
Db	381	GGAGTCCCTCCCATCTCTCTGAGCCAGGGAGTAGGAGTGGTATTCCGACGGGATGG	322							
Qy	1709	GGGTGTCTTGGGGGTGGAGATAGGGGAGCAGAGGAGCTATTGTCTAAGGCCCGATAGG	1768							
Db	321	GGGTGTCTTGGGGGTGGAGATGGGGGAGCAGAGGAGCGATTGCTAAGGCCCAATAGG	262							
Qy	1769	CACCTCATTTGCCGGGAATGTGCCCAGGGAGCAGTGGTGGTTATAACTCAGGCCCGGT	1828							
Db	261	CACCTCATTTGCCGGGAATGTGCCCAGGGAGCAGTGGTGGTTATAACTCAGGCCCTGGT	202							
Qy	1829	GCCCAGAGCCCGAGGAGGCGAGTGGCCCGAGAAGGCCACAGGCCCTGAGAAGTCTCGGCTG	1888							
Db	201	GCCCAGAGCCCGAGGAGGCGAGTGGCCCGAGAAGGTGTCAGGCCCTGAGAAATCCCGGCTG	142							
Qy	1889	AGCTGGAGGCAAAATCCCCACCCCTTACCTTGGGGGACAGGGCAAGTACACCTGTGAGG	1948							
Db	141	AACTGGAAGCGAATCCCCACCCCTTACCTTGGGGGACAGGGCAGTGTGAGACCTGGGAGG	82							
Qy	1949	GTGCTCAGCAGGCGAGGAGGAGTGTCTGTGCGTCTCTGCACCCACATCTTTTCTC-T	2007							
Db	81	GTGCTCAGCAGGCGAGGAGGAGTGTCTGTGCGTCTCTGCACCCACATCTTTTCTCTT	22							
Qy	2008	GTCCCCCTCTTGGCCCTGTCTG 2028								



	/db_xref="RZPD:DKFzp686N01223Q"
	/db_xref="taxon:9606"
	/clone="DKFzp686N01223"
	/tissue_type="small intestine"
	/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiI + SfiII"
	/dev_stage="adult"
	/note="hypothetical protein, N-terminus truncated, not fully spliced"
gene	1..5511
	/gene="DKFzp686N01223"
CDS	<2191..2811
	/gene="DKFzp686N01223"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="CAH10564.1"
	/db_xref="GI:50949597"
	/translation="ITGSHFLMLNSTPLCGSLCQTIVVGVSVILYSSRACYNLWV RLISQDTLPAGMINSYSRRAYFFDNPRIYDDSDLLPRLGSREGSLPSQSGLGWGT MTGCGSSYVTVPHLNGPMDTAPLLFTCSNLDLNNHHSLYVTPQN"
ORIGIN	
Query Match	6.5%; Score 324; DB 3; Length 5511;
Best Local Similarity	71.1%; Pred. No. 6e-19;
Matches 513; Conservative	0; Mismatches 175; Indels 34; Gaps 5;
QY _	355 TTTT-----TTTTTTTTTGACGGAGTCTCGCTCTGCGCCAGGCTGGA-GTGCA 413
Db	1787 TATCTTTTTTTTTTTTTTTTGAGGTGGAGTCTCACTGTGCCCCCAGGCTGGAGGTGCA 1728
QY	414 GTGCGGGGATCTCGGCTCACTGCAGCTCCGCGCTCCGCGGTTCACGCCATTCTCTCGCT 473
Db	1727 GTGGCATGACTCACTGCTCAAGCTCCGCTCTCGGCTCTGGGCTCACCGTCTCTCTGCC 1668
QY	474 CAGCCTCCCCAAGTAGCTGGGACTACAGCGCCCGCCACTACGCGCGGCTAAATTTTTTGT 533
Db	1667 CAGCCTCCCGATTAGCCGGACTACAGCACCGCCACCAGCGCGGC-AAATTTTTTGT 1609
QY	534 ATTTTTAGTAGAGACGGGTTTACCGTTTTAGCGGGATGGCTCGATCTCTGACCTC 593
Db	1608 ATTTTTAGTAGAGACAGGGTTTACCGGTTTAGCGGATGGTCTCGATCTCTGACCTC 1549
QY	594 GTGATCCGCGCGCTCGGCTCCCAAAGTGTCTGGATT---ACAGCGGTGAGCCACCGCG 650
Db	1548 GTGATCCACCGCGCTTGGCTCTCCAGCGCTGGGATTACACAGCGCTGAGCCACTGCG 1489
QY	651 CCCGGCCATGATCATCTCTT-----GACTATGCTGATGTGACAA 690
Db	1488 CTGGCCCAAAAACATTTATATATACCCCATACATATACCTTATAATCACCCACAAAA 1429
QY	691 GTACTTAAGCCATCAGACTCTACCTTTAAATATGAGTTTGGCCAGGACCGTGGCT 750
Db	1428 TTAAAAATAAAAAATTTTCTAAGAATTAaaaaaaataatATATGGCTGAGTACCATGGCT 1369
QY	751 CATGCTGTAATTCAGCACTTTGGGAGGACAGGTGGTGAATCACTTAGGCGCAGAG 810
Db	1368 CATGCTTGCAATCCAGCACTTTTGAAGGCCAAGATACAGGATCATCTGAGGCCAGAG 1309
QY	811 TTTGAGACCCAGCTGGCCAAACATGGTGAAACTCTGTCTTTACTAAAAAaaaaaaaa 870
Db	1308 TTCAGAGCAGCTGGGCCAACATAGCAAGACTTGTCTCTCCAAGAAAGAAAAAAAAA 1249
QY	871 AAAAAAATCAGCGGGTGTCTGGGGACACCTGTAATCCAGCTATCTGTGAGGCTGA 930
Db	1248 A-----ASCTGAGCGTGTGTGTAATGCTCGCACTTCCAGCCACTGGGAGGCTGA 1198
QY	931 GGCAAGAGTCACTTTGAACCTTGAGCGGAGGTTGAGTGGCCGAGATCACATCACC 990
Db	1197 GGCAGGAGAAATTAATAAGCCAGAGTTTGAGGTTACAGTAAGCTATAATTGGCCACT 1138
QY	991 GCCCTCCAGCTGGCGCAGAGCAAGACTCTGTCTCAATTAATAATAAACAAACGAA 1050

Db	1137	ACATTCAGCCTGGGTGACACAGCTGAGATGTCTCCATCAAGAAAAAAGTAATGCA	1076
Qy	1051	CA 1052	
Db	1077	AA 1076	
RESULT 14			
LOCUS	HSM807092	6934 bp	mRNA
DEFINITION	Homo sapiens mRNA; cDNA DKFP686O10123 (from clone DKFP686O10123).		linear
ACCESSION	BX640943		
VERSION	BX640943.1	GI:34365276	
KEYWORDS	HTC		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 6934) Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.		
CONSTRIM	The German cDNA Consortium		
TITLE	Direct Submission		
JOURNAL	Submitted (13-JUL-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY		
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Oigen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFP686O10123) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: <a href="http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFP686O10123">http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFP686O10123</a> Further information about the clone and the sequencing project is available at <a href="http://mips.gsf.de/projects/cdna/">http://mips.gsf.de/projects/cdna/</a> .		
FEATURES	Location/Qualifiers		
source	1..6934		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DKFP686O10123"		
	/tissue_type="uterus"		
	/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfilA + SfiIB"		
	/dev_stage="adult"		
	/note="genomic"		
ORIGIN			
Query Match	6.5%;	Score 322.8;	DB 3; Length 6934;
Best Local Similarity	70.5%;	Pred. No. 6.5e-19;	
Matches	530;	Conservative 0;	Mismatches 162; Indels 60; Gaps 5;
Qy	346	TC TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGGAGTCTCGCTCTGTGCGCCAGGCT	405
Db	5398	TTTTTGTTTTGTTTTGTTTTGTGTTTCTTGAGACAGTCTCACTCTGTGTGCCAGGCT	5457
Qy	406	GGAGTCAGTGGCGGGATCTCGGCTCACTGCAAGCTCGGCTCCCGGTTTCACGCCATTC	465
Db	5458	GGAGTCAGTGGCATGATATCTGCACACTGCAGCCCTCTCTTCCGGGGTCAGATATTC	5517
Qy	466	TCCTGCTCTAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACGCCCGGCTAAT	525
Db	5518	TCCAGCCTCCGCTCTTGTAGTAGCTGGATTGGAGGCGCATGCTACCACGCCAGCTAA-	5576
Qy	526	TTTTTTGTATTTTGTAGTAGACCGGGGTTTCAACGTTTTAGCCGGGATGGGCTCGATCTC	585
Db	5577	-TTTTTGTATTTTGTGTAGACACCGGATTTCCACATGTGGCCAGGCTGGTCTTTGAGCTT	5633
Qy	586	CTGACCTC--GTGATCCGCCCGCTCTCGGCTCCCAAAAGTCTGGGATTTACAGGCGTGAGC	643
Db	5636	CTGACCTCAGTGAATTCGCTGCGCTTCGCGCTTCCCAAGTCTGGGATTTGAGGCGTGAGG	5695
Qy	644	CACGCGCCCGGGCC-----ATGATCATCTTCT	670





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 2885.07 Seconds

(without alignments)  
16811.954 Million cell updates/sec

**Title:** US-09-936-271C-13 COPY 6000 7000

Perfect score:

Sequence: 1 gacaaccaagagccccaag.....gttactgtattggatggcac 1001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs. 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

```

Post-processing: Minimum March 0%

100% Processing: Minimum Match 0%  
Maximum Match 100%

Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl : \*

1: qb ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om: \*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gfb\_pr:\*

```
10: gb_ro:*
```

II: gb\_sts:

12: gb\_sy: \*

```
13: gb_un: *
14: gb_un: *
```

```
l4: gb_v1:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1001	100.0	11570	6	CQ788219	Sequence
	1001	100.0	11570	6	CQ874895	Sequence
	1001	100.0	11570	6	CQ874861	Sequence
	1001	100.0	11570	9	AF135028	Homo sapi
C 5	1001	100.0	107487	9	AC011483	Homo sapi
	1001	100.0	107487	9	AF243527	Homo sapi
C 6	984.2	98.3	217346	9	AC027602	Homo sapi
C 7	984.2	98.3	217346	9	AC130782	Homo sapi
C 8	958.8	95.8	200792	2	AC098800	Homo sapi
C 9	284.6	28.4	142003	9	AC133961	Homo sapi
C 10	276.8	27.7	177787	9	AC133961	Homo sapi
C 11	274.4	27.4	138411	9	HSJ071H19	Human DNA
C 12	262.4	26.2	92882	9	AL355815	Human DNA
C 13	261.6	26.1	85304	9	HS22715	Human DNA
C 14	260.26	26.0	161014	9	AC087591	Homo sapi
C 15	260.26	26.0	177876	2	AC025163	Homo sapi
C 16	256.4	25.6	177562	9	HS269WM15	Human DNA
C 17	253.6	25.3	136649	9	HSJ0635E8	Human DNA
C 18	253.2	25.3	173840	9	AC092445	Homo sapi
C 19	253.2	25.3	206819	2	AC016029	Homo sapi

## ALIGNMENTS

RESULT 1  
 LOCUS CQ788219 11570 bp DNA linear PAT 24-MAR-2004  
 DEFINITION Sequence 2 from Patent WO2004021008.  
 ACCESSION CQ788219  
 VERSION CQ788219.1 GI:45723068  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Diamandis, E.P.  
 TITLE Methods for detecting breast and ovarian cancer  
 JOURNAL Patent: WO 2004021008-A 2 11-MAR-2004;  
 MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)  
 FEATURES  
 Location/Qualifiers  
 1..11570  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 100.0%; Score 1001; DB 6; Length 11570;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-228;  
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GACAACCAAGAGCCCAAGGTGAGTGTCCAGGTCTCTTGATACCGACCATCTCTGC 60  
 Db 6000 GACACCAAGAGCCCAAGGTGAGTGTCCAGGTCTCTTGATACCGACCATCTCTGC 6059  
 QY 61 CGCCTTCCATCTTTCTCCACTTCTCATGTGTTCTGTTGACAGTGCATTCCTTAAGG 120  
 Db 6060 CGCCTTCCATCTTTCTCCACTTCTCATGTGTTCTGTTGACAGTGCATTCCTTAAGG 6119  
 QY 121 TCCTCCAGTCTTGAATATCAGCGTGCTAAGTCAGAAAGGTCGAGGATGCTTACCCGA 180  
 Db 6120 TCCTCCAGTCTTGAATATCAGCGTGCTAAGTCAGAAAGGTCGAGGATGCTTACCCGA 6179  
 QY 181 GACAGATAGATGACCAATGTCTTCGCCCGGTGACAAAGCAGGTAGAGACTCCTGCCAGG 240  
 Db 6180 GACAGATAGATGACCAATGTCTTCGCCCGGTGACAAAGCAGGTAGAGACTCCTGCCAGG 6239

		/mol_type="unassigned DNA" /db_xref="taxon:9606"	
ORIGIN			
Query Match		100.0%; Score 1001; DB 6; Length 11570;	
Best Local Similarity		100.0%; Pred. No. 9,7e-228;	
Matches 1001; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	GACAAACCAAGAGCCCCCAAGTGAGTGTCCAGGTTCTTCTTGATACCGACCCATCTCTGC	60
Db	6000	GACAAACCAAGAGCCCCCAAGTGAGTGTCCAGGTTCTTCTTGATACCGACCCATCTCTGC	6059
Qy	61	CGCTTCCATCTTTCTCCACTTCTCATTTGTCTTCTGTTGACAGTGACATTCCTTAAG	120
Db	6060	CGCTTCCATCTTTCTCCACTTCTCATTTGTCTTCTGTTGACAGTGACATTCCTTAAG	6119
Qy	121	TCCTTCAGTGTGAAATATCAGCGTGTCTTAAGTGCAGAAAAGTGCAGAGTGCCTTACCCGA	180
Db	6120	TCCTTCAGTGTGAAATATCAGCGTGTCTTAAGTGCAGAAAAGTGCAGAGTGCCTTACCCGA	6179
Qy	181	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTTGCCAGG	240
Db	6180	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTTGCCAGG	6239
Qy	241	TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCCTCACTCGGTAACTGGAGC	300
Db	6240	TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCCTCACTCGGTAACTGGAGC	6299
Qy	301	GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGAGA	360
Db	6300	GTTGCCAAATCTGAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTTCTCAGAGA	6359
Qy	361	GCTCATACCTTAGAGTAGTGTGTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC	420
Db	6360	GCTCATACCTTAGAGTAGTGTGTTAGTGAATAATGCTGAGCTGCTTATGTCATTTC	6419
Qy	421	CAGTTTTTTAGTAGCCACATTAATAACAGGTAAAAAGGCTGGGGCAGTGGCTCACACT	480
Db	6420	CAGTTTTTTAGTAGCCACATTAATAACAGGTAAAAAGGCTGGGGCAGTGGCTCACACT	6479
Qy	481	GTAATCCAGACATTTTGGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGAGTTTGAGA	540
Db	6480	GTAATCCAGACATTTTGGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGAGTTTGAGA	6539
Qy	541	CTAGCCTGGCCAAACATGCGGAACTCTGCTCTCTAAAAAATAACAAAAATTTAGCTCGC	600
Db	6540	CTAGCCTGGCCAAACATGCGGAACTCTGCTCTCTAAAAAATAACAAAAATTTAGCTCGC	6599
Qy	601	ATGTTGGCGGGCGCTGTAAATCTCAGCTGTCTCAGAGGCCGAGACACAAAGATCACTTAA	660
Db	6600	ATGTTGGCGGGCGCTGTAAATCTCAGCTGTCTCAGAGGCCGAGACACAAAGATCACTTAA	6659
Qy	661	ACCAGAGGTGGAGGTTGAGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	720
Db	6660	ACCAGAGGTGGAGGTTGAGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	6719
Qy	721	AGAGTGACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG	780
Db	6720	AGAGTGACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG	6779
Qy	781	TTAACTTTTAATACCCCAATGTATCCCAATACAAATCAATTTCAAAGTGTAAATATAA	840
Db	6780	TTAACTTTTAATACCCCAATGTATCCCAATACAAATCAATTTCAAAGTGTAAATATAA	6839
Qy	841	ACAATATGAAATGAGATACCTTTAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	900
Db	6840	ACAATATGAAATGAGATACCTTTAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	6899
Qy	901	GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC	960
Db	6900	GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC	6959
Qy	961	AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC	1001

RESULT 2  
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 3 from Patent WO2004077060.  
ACCESSION CQ874885  
VERSION CQ874885.1 GI:52748035

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1  
AUTHORS Diamandis, E.P. and Petraki, C.  
TITLE Assay for detection of renal cell carcinoma  
JOURNAL Patent: WO 2004077060-A.3 10-SEP-2004;  
MOUNT Mount Sinai Hospital (CA)  
FEATURES  
1..11570  
Location/Qualifiers  
/organism="Homo sapiens"

Db	6960	AGTAGCCACATGGCTAGCAGTACTGTTATTGGATGGCAC	7000
RESULT 3	CQ874961	CQ874961	PAT 27-SEP-2004
LOCUS	Sequence 5 from Patent WO2004075713.	11570 bp	DNA linear
DEFINITION	CQ874961		
ACCESSION	CQ874961		
VERSION	CQ874961.1	GI:52748060	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1	Diamandis,E.P. Multiple marker assay for detection of ovarian cancer Patent: WO 2004075713-A 5 10-SEP-2004;	
AUTHORS	MOUNT SINAI HOSPITAL CORPORATION (CA)		
TITLE	Location/Qualifiers		
JOURNAL	1. .11570		
FEATURES	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
source			
ORIGIN			
Query Match	100.0%; Score 1001; DB 6; Length 11570;		
Best Local Similarity	100.0%; Pred. No. 9, 7e-28;		
Matches 1001; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1	GACAACAAAGAGCCCCAAGGTGAGTGTCAGAGTCTTCTTGATACCAGCATCTCTGC	60
Db	6000	GACAACAAAGAGCCCCAAGGTGAGTGTCAGAGTCTTCTTGATACCAGCATCTCTGC	6059
Qy	61	CGCTTCCATCTTCTCCACTTCTCATTTGTTCTCTGTTTGACATGCACTTCCCTAAGG	120
Db	6060	CGCTTCCATCTTCTCCACTTCTCATTTGTTCTCTGTTTGACATGCACTTCCCTAAGG	6119
Qy	121	TCCTCCAGTCTTGAATATCAGGCTGCTAAAGTCAGAAAAGGTCGAGGATGCTTACCCGA	180
Db	6120	TCCTCCAGTCTTGAATATCAGGCTGCTAAAGTCAGAAAAGGTCGAGGATGCTTACCCGA	6179
Qy	181	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCTGCCAGG	240
Db	6180	GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCTGCCAGG	6239
Qy	241	TGAGGACACCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	300
Db	6240	TGAGGACACCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	6299
Qy	301	GTTCGCCAAATCTGAGAATCCAGCAATTCGCAAGACAGTCAGGACCCCTGTTCTCACAGA	360
Db	6300	GTTCGCCAAATCTGAGAATCCAGCAATTCGCAAGACAGTCAGGACCCCTGTTCTCACAGA	6359
Qy	361	GCTCATACCTCAGAGTAGTGGTGTGTTAGTAGAAATAAGTCTGAGCTGCTATGTCATTTC	420
Db	6360	GCTCATACCTCAGAGTAGTGGTGTGTTAGTAGAAATAAGTCTGAGCTGCTATGTCATTTC	6419
Qy	421	CAGTTTTTTAGTAGCCACATTAACACAGGTAAACAAAGGCTGGGCGCAGTGGCTCACACT	480
Db	6420	CAGTTTTTTAGTAGCCACATTAACACAGGTAAACAAAGGCTGGGCGCAGTGGCTCACACT	6479
Qy	481	GTAATCCAGCAGCTTTGGGAGGCTGAGGAGGAGCAGATCACCTTTGCTCAGGAGTTTGAGA	540
Db	6480	GTAATCCAGCAGCTTTGGGAGGCTGAGGAGGAGCAGATCACCTTTGCTCAGGAGTTTGAGA	6539
Qy	541	CTAGCTGGCCAACTGGCGAAACTCTGCTCTCTAAAAAATAACAAAAATTTAGCTGGC	600
Db	6540	CTAGCTGGCCAACTGGCGAAACTCTGCTCTCTAAAAAATAACAAAAATTTAGCTGGC	6599
Qy	601	ATGCTGGCGGCGCTGTAACTTCTAGCTGCTCAGGAGGCCGAGACACAAGAAATCACTTAA	660

Db	6600	ATGGTGGCGGCGCTGTAACTTCTAGCTGCTCAGGAGGCCGAGACACAAGAATCACTTAA	6659
Qy	661	ACCCAGGAGTGGAGGTTGCACTGAGCTGAGATCGTCCACTCTACTCCAACCTGGGAGAC	720
Db	6660	ACCCAGGAGTGGAGGTTGCACTGAGCTGAGATCGTCCACTCTACTCCAACCTGGGAGAC	6719
Qy	721	AGAGTGACACTTTTCTCTCAAAGAAAAAAGAAAAAAGTAAAGAAAAAGAACAGGTGAAG	780
Db	6720	AGAGTGACACTTTTCTCTCAAAGAAAAAAGAAAAAAGTAAAGAAAAAGAACAGGTGAAG	6779
Qy	781	TTAACTTTTAAATAACCAATGTATCCCAATCAATCAATTTTCAAAAGTGTAAATTAATATAA	840
Db	6780	TTAACTTTTAAATAACCAATGTATCCCAATCAATCAATTTTCAAAAGTGTAAATTAATATAA	6839
Qy	841	ACAAATATGATGAGATCTTTTACATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	900
Db	6840	ACAAATATGATGAGATCTTTTACATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	6899
Qy	901	GAGTATATATGTTTCTGACAGCAGCATCTCAATTTGGAGCTAGCTACATTTTCTGAGTCTC	960
Db	6900	GAGTATATATGTTTCTGACAGCAGCATCTCAATTTGGAGCTAGCTACATTTTCTGAGTCTC	6959
Qy	961	AGTAGCCACATGGCTAGCAGTACTGTTTGGATGGCAC	1001
Db	6960	AGTAGCCACATGGCTAGCAGTACTGTTTGGATGGCAC	7000
RESULT 4	AF135028	AF135028	PRI 27-JUN-2000
LOCUS	Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.	11570 bp	DNA linear
DEFINITION	AF135028		
ACCESSION	AF135028.1	GI:4589282	
VERSION			
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1	(bases 1 to 11570) The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation J. Biol. Chem. 274 (53), 37511-37516 (1999)	
AUTHORS	Yousef,G.M. and Diamandis,E.P.		
TITLE	The new human kallikrein gene family: implications in carcinogenesis		
JOURNAL	Trends Endocrinol. Metab. 11 (2), 54-60 (2000)		
MEDLINE	21121728		
PUBMED	10675891		
REFERENCE	3	(bases 1 to 11570) Yousef,G.M., Luo,L.Y. and Diamandis,E.P. Direct Submission Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada	
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES	Location/Qualifiers		
source	1. .11570 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /map="19q13.3-q13.4" Join(2101..2131,2210..2293,4762..5023,5763..6019, 6105..6238,11092..11570) /product="kallikrein-like protein 2 KLK-L2" 		

/db\_xref="GI:4589283"  
/translation="MATARPPMMWVLICALITALLLVTEHVLANNVSDHPNSNTVPS  
GSDNLGAGAGEDARSDDSSRIINGSDCMHTOPWQALLLRPNOLYCGAVLVHPOM  
LLTAHCKKKVVRVLGHYSLPVYESQOMFQVKSI PHPGYSHPGSHNDJMLIKLN  
RRIRPTKDRVINVSCHPSAGTKCLVSGWGTITKSPQVHPFKVLQCLNLSVLSQKRC  
DAYPRQIDDDTFMFCAGDKAGRDSQGDSPVVCNGSLVSGWGDYPCARPNRPGVYT  
NLCKFTKWIQTIQANS"

ORIGIN

Query Match 100.0%; Score 1001; DB 9; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 9,7e-228;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAAACAGAGCCCCCAAGGTGAGTCCAGGTCTCTTGTATACCGACCCATCTCTGC 60  
Db 6000 GACAAACAGAGCCCCCAAGGTGAGTCCAGGTCTCTTGTATACCGACCCATCTCTGC 6059

Qy 61 CGCCTTCATCTTTCTCCACTCTCTCATTTGTTCTCTGTTGACAGTGCACTTCCTTAAG 120  
Db 6060 CGCCTTCATCTTTCTCCACTCTCTCATTTGTTCTCTGTTGACAGTGCACTTCCTTAAG 6119

Qy 121 TCCTCCAGTGCTTGAATATCAGCTGCTAACTCAGAAAGTGCAGAGTCTTACCGA 180  
Db 6120 TCCTCCAGTGCTTGAATATCAGCTGCTAACTCAGAAAGTGCAGAGTCTTACCGA 6179

Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGG 240  
Db 6180 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGG 6239

Qy 241 TGAGGACACCTCTCTTTATCAGCAGATACACACTGAGTGCGCAACTCGGTAACATGAGC 300  
Db 6240 TGAGGACACCTCTCTTTATCAGCAGATACACACTGAGTGCGCAACTCGGTAACATGAGC 6299

Qy 301 GTTGCCAAATCTGAGATTCAGCAATTCGCAAGCAGTCAAGACCCCTGTTCTCACAGA 360  
Db 6300 GTTGCCAAATCTGAGATTCAGCAATTCGCAAGCAGTCAAGACCCCTGTTCTCACAGA 6359

Qy 361 GCTCATACCTTAGAGTAGTGTGTTAGTGAATAATGCTGAGTGCTTATGTCATTTTC 420  
Db 6360 GCTCATACCTTAGAGTAGTGTGTTAGTGAATAATGCTGAGTGCTTATGTCATTTTC 6419

Qy 421 CAGTTTTTTTAGTAGCCACATTTAAACAGGTAAAAAGGCTGGCGCAGTGCTCACACCT 480  
Db 6420 CAGTTTTTTTAGTAGCCACATTTAAACAGGTAAAAAGGCTGGCGCAGTGCTCACACCT 6479

Qy 481 GTAATCCAGCACTTTGGAGGCTGAGCAGCAGATCACCTTTGGTCAGAGTTTGA 540  
Db 6480 GTAATCCAGCACTTTGGAGGCTGAGCAGCAGATCACCTTTGGTCAGAGTTTGA 6539

Qy 541 CTAGCCTGGCCAAACATGGCGAACTCTCTCTTAAAAAAAATAACAAAAATTTAGCCTGGC 600  
Db 6540 CTAGCCTGGCCAAACATGGCGAACTCTCTCTTAAAAAAAATAACAAAAATTTAGCCTGGC 6599

Qy 601 ATGTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAAGATCACTTAA 660  
Db 6600 ATGTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAAGATCACTTAA 6659

Qy 661 ACCCAGGAGGTGGAGTTGCGAGTGAGATCGTGCACTCACTCAACCTGGGAGAC 720  
Db 6660 ACCCAGGAGGTGGAGTTGCGAGTGAGATCGTGCACTCACTCAACCTGGGAGAC 6719

Qy 721 AGAGTGACACTTTTCTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG 780  
Db 6720 AGAGTGACACTTTTCTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAG 6779

Qy 781 TTAACCTTTAATACCAATGATATCCAAATACAATCAATTTCAAAGTGTAATTAATAA 840  
Db 6780 TTAACCTTTAATACCAATGATATCCAAATACAATCAATTTCAAAGTGTAATTAATAA 6839

Qy 841 ACAATTATGAATGACATCTTACATTTCTTTCTTTTTCATATTAAGTCTTTGAAAGT 900  
Db 6840 ACAATTATGAATGACATCTTACATTTCTTTCTTTTTCATATTAAGTCTTTGAAAGT 6899

Qy 901 GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGCTC 960  
Db 6900 GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGCTC 6959

Qy 961 AGTAGCCACATGTGGCTAGCAGTACTGTATTTGGATGGCAC 1001  
Db 6960 AGTAGCCACATGTGGCTAGCAGTACTGTATTTGGATGGCAC 7000

RESULT 5  
AC011483/c  
LOCUS  
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.  
AC011483  
AC011483.7 GI:21637461  
VERSION  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 107487)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Unpublished  
JOURNAL 2 (bases 1 to 107487)  
REFERENCE Direct Submission  
AUTHORS Direct Submission  
TITLE Direct Submission  
JOURNAL 3 (bases 1 to 107487)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE 1 (bases 1 to 107487)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
TITLE Direct Submission  
JOURNAL 1 (bases 1 to 107487)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.6.  
NOTE: Shatter libraries failed to resolve dinucleotide repeat.  
Unsure number of repeat copies 64998-65494. Forced join 65015.  
Location/Qualifiers  
1. 107487  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clones="CTB-147C22"  
64998..65494  
/note="NOTE: Shatter libraries failed to resolve  
dinucleotide repeat. Unsure number of repeat copies  
64998-65494. Forced join 65015."

misc\_feature  
ORIGIN  
Query Match 100.0%; Score 1001; DB 9; Length 107487;  
Best Local Similarity 100.0%; Pred. No. 1.2e-227;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAAACAGAGCCCCCAAGGTGAGTGTCCAGGTCTCTTGTATACCGACCCATCTCTGC 60  
Db 54220 GACAAACAGAGCCCCCAAGGTGAGTGTCCAGGTCTCTTGTATACCGACCCATCTCTGC 54161

Qy 61 CGCCTTCATCTTTCTCCACTCTCTCATTTGTTCTGTTGACAGTGCACTTCCTTAAG 120  
Db 54160 CGCCTTCATCTTTCTCCACTCTCTCATTTGTTCTGTTGACAGTGCACTTCCTTAAG 54101

Qy 121 TCCTCCAGTGCTTGAATATCAGCTGCTAACTCAGAAAGTGCAGAGTCTTACCCGA 180  
Db 54100 TCCTCCAGTGCTTGAATATCAGCTGCTAACTCAGAAAGTGCAGAGTCTTACCCGA 54041

Qy 181 GACAGATAGATGACACCATGTTCTTGGCGCGGTGACAAAGCAGGTAGAGACTCTCTGCCAGG 240

```
Db 54040 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCCTGCCAGG 53981
Qy 241 TGAGGACACCTCTCTTTATTGAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC 300
Db 53980 TGAGGACACCTCTCTTTATTGAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC 53921
Qy 301 GTTGCCAAATCTGAGAACTCAGCAATGCGCAAGACAGTCAAGGCCCTGTTCTCACAGA 360
Db 53920 GTTGCCAAATCTGAGAACTCAGCAATGCGCAAGACAGTCAAGGCCCTGTTCTCACAGA 53861
Qy 361 GCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGAGCTGCTTATGTCATTTTC 420
Db 53860 GCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGAGCTGCTTATGTCATTTTC 53801
Qy 421 CAGTTTTTTAGTAGCCACATTAAACAGGTAAAAAGGCTGGGCGAGCTGCTCACACCT 480
Db 53800 CAGTTTTTTAGTAGCCACATTAAACAGGTAAAAAGGCTGGGCGAGCTGCTCACACCT 53741
Qy 481 GTAATCCAGCACCTTGGGAGGCTGAGCGAGCAGATCACTTTGGTCAGGAGTTTGAGA 540
Db 53740 GTAATCCAGCACCTTGGGAGGCTGAGCGAGCAGATCACTTTGGTCAGGAGTTTGAGA 53681
Qy 541 CTAGCTGCGCAACATGCGCAAACTCTCTCTTAAAAAATAACAAAAATTAGCCTGGC 600
Db 53680 CTAGCTGCGCAACATGCGCAAACTCTCTCTTAAAAAATAACAAAAATTAGCCTGGC 53621
Qy 601 ATGTTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAAGAACTCACTTAA 660
Db 53620 ATGTTGGCGGGCGCTGTAATCTCAGCTGCTCAGGAGCCGAGACACAAAGAACTCACTTAA 53561
Qy 661 ACCGAGGAGTGGAGTTGCGAGTGAAGTCTGCTCAAAAGAAAAAACAAGTAAAAAGAACAGTGAAG 780
Db 53560 ACCGAGGAGTGGAGTTGCGAGTGAAGTCTGCTCAAAAGAAAAAACAAGTAAAAAGAACAGTGAAG 53501
Qy 781 AGAGTGACACTTTTGTCTCAAAAGAAAAAACAAGTAAAAAGAACAGTGAAG 780
Db 53500 AGAGTGACACTTTTGTCTCAAAAGAAAAAACAAGTAAAAAGAACAGTGAAG 53441
Qy 781 TTAACCTTTAATACCCCAATGATCCCAATCAATCAATTTCAAAAGTGAATTAATATAA 840
Db 53440 TTAACCTTTAATACCCCAATGATCCCAATCAATCAATTTCAAAAGTGAATTAATATAA 53381
Qy 841 ACAATTATGAATGAGATCTTACATCTCTTTCTTTGTTTCAATTAAGTCTTTGAAAGT 900
Db 53380 ACAATTATGAATGAGATCTTACATCTCTTTCTTTGTTTCAATTAAGTCTTTGAAAGT 53321
Qy 901 GAGTATATGTTATGCTGAGACACATCTCAATTTGGACTAGCTACATTTTCAGTGTCTC 960
Db 53320 GAGTATATGTTATGCTGAGACACATCTCAATTTGGACTAGCTACATTTTCAGTGTCTC 53261
Qy 961 AGTAGCCACATGTGGCTAGCAGTACTCTGATTGGATGGCAC 1001
Db 53260 AGTAGCCACATGTGGCTAGCAGTACTCTGATTGGATGGCAC 53220
```

## RESULT 6

AF243527/c  
LOCUS AF243527 230000 bp DNA linear PRI 21-NOV-2000  
DEFINITION Homo sapiens serine protease gene cluster, complete sequence.

ACCESSION AF243527

VERSION AF243527.1 GI:11244757

## KEYWORDS

SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,

Moss, P., Paepel, B. and Wang, K.

TITLE Sequencing and expression analysis of the serine protease gene

cluster located in chromosome 19q13 region

## JOURNAL

Gene 257 (1), 119-130 (2000)

MEDLINE  
PUBMED

20510030  
11054574

2 (bases 1 to 230000)

## REFERENCE

## AUTHORS

Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,

Moss, P., Paepel, B. and Wang, K.

## TITLE

Direct Submission

Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,

Bothell, WA 98021, USA

## FEATURES

Location/Qualifiers

1..230000

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="19"

/map="19q13"

complement(<6963..>11517)

/gene="KLK1"

complement(join(<6963..7118,7668..7804,7923..8212,

9482..9641,11472..>11517))

/gene="KLK1"

/product="renal kallikrein"

complement(join(6963..7118,7668..7804,7923..8212,

9482..9641,11472..11517))

/gene="KLK1"

/note="serine protease"

/codon\_start=1

/product="renal kallikrein"

/protein\_id="AAG33353.1"

/db\_xref="GI:11244758"

/translation="MWPLVCLALSLGGTGAAPFIQSRIVGVGMECEHSPQWQALVH

FTSFCGGILVHRQWVLTAAHCTSDNYQLMGRHNLFDDETAFVHVSSEFPHPGEN

MSLLENHTRQADEYSHDLMLSLTEPADTITDAVKVELPTEPEHSGTCLASGWS

IEBENSFPDDLQCVLDKILPNBECKKAHVQVDFMLCVLHSGGKDTVCVDSGGPL

MCQGVQGVTSWGVPCGTGTPKPSVAVRVLSVKWIEDTIAENS"

complement(join(<13552..13704,14377..14530,14678..14917,

15416..15569,19204..>19246))

/product="ACO protease"

complement(join(13552..13704,14377..14530,14678..14917,

15416..15569,19204..19246))

/note="serine protease"

/codon\_start=1

/product="ACO protease"

/protein\_id="AAG33354.1"

/db\_xref="GI:11244759"

/translation="MWLLTSLFLLASTAAQDGKLLLEGDECAHSPQWQALVHGR

FNGCASLISPHWLSAAHCSRFMRVRLGHEHLKRDGPQLRTTSRVIPHPYEARS

HRNDIMLRVQPARLNQVPRPAVLPTRCPPHGEACVVGWGLVPLSPVSLPDLHC

ANISLISDTSCKDSYPCRLTNTMVCAGAEGRGAESCEGDSGGPLVCGGILQGVSWGD

VPCDNTKPGVYTKVCHYLEWIRETKEN"

<42595..>47769

/gene="KLK3"

join(<42595..42640,43880..44039,45669..45955,46099..46235,

47614..>47769)

/gene="KLK3"

/product="prostate specific antigen"

join(42595..42640,43880..44039,45669..45955,46099..46235,

47614..47769)

/gene="KLK3"

/note="serine protease"

/codon\_start=1

/product="prostate specific antigen"

/protein\_id="AAG33355.1"

/db\_xref="GI:11244760"

/translation="MWVTVVFLTSLVTWIGAAPLILSRIVGVGMECEKHSQWQVLVAS

RGRAVCGGVLPQWVLTAAHCTRNKSVILLGRSLFHPEDTGVQVFOVSHSFPHPLYD

MSLLKNFLRGDDSSHDMLRLSEPAELTDVAVMDLPTQEPALGTTTCVASCWSGI

EPFELTPKKLQCVLDLHIVSNDVCAQVHPQKVTKFMLCAGRWTCGKSTCSDSGSGLV

CNGVLQGITSWGSEPCALPERPSLYTKVHYRWIKDTIVANP"

<61139..>66229

/gene="KLK2"

join(<61139..61184,62391..62550,64142..64428,64542..64678,

66074..>66229)

/gene="KLK2"

```
CDS
/product="glandular kallikrein"
/join(61139..61184,62391..62550,64142..64428,64542..64678,66074..66229)
/gene="KLK2"
/notes="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MWDLVLSIALSVGTGAVPLIQSRIVGGVCEKEHSOPWQVAVYS
HGMACHGGLVHPQWLTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
SLKLHQSLRDESDHDLMLRLSEPAKITDVVKVLPTQEPALGTTCTVASGWSI
EPSEFLPRSLQCVSLHLSNDMCARAYSEKVFMLCAGLWGTGTCGDSGGPLV
CNGVLOGTSNGPPFCALPEKPAVYTVVHYRWIKDTIAANP"
/complement(join(<94602..94754,96027..96163,96247..96497,96919..97081,98345..98405))
/product="serine protease"
/complement(join(94602..94754,96027..96163,96247..96497,96919..97081,98345..98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPGWFLGVLILVAGSLVSGSCQIINGEDCSPHSQPW
QAALVMNELFCSGVLVHPQWLTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
VRHPEYRPLANDMLIKLDESVESEDITRSIASQCPTAGNSCLVSGWGLIANGR
MPTVLQCVNVSVEEVCSLYDPLHYPSMFCAGGHDQKDCNCGDSGGPLICNGYLQ
GLVSFGKAPCGQGVPGVYTNLCKFTIEWIKTVQAS"
/complement(<113101..>140303)
/gene="KLK5"
/notes="synonym: SCTE"
/complement(join(<113101..131456,136310..136443,136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
/complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translation="MWLWLCITALLLGVTDAKSDSSRIINGSDDCMHTQPWQAAAL
LLRNQYCGAVLHPQWLTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
PGYHPGHSNDMLIKLNRIRPTKDVPIVNSHCPKAGTCKLVSQWGTCKSPQVHF
PKVLQCNISVLQKCEDAYPRQIDTFCAGKAGKAGDSQCGSGSPVVCNGLSLQSL
VSGDYPACNRFQVYNLCKFTKWIQTIQANS"
/complement(join(<146834..146986,149628..149764,151186..151433,155052..155208,155948..>155987))
/product="protease M"
/complement(join(146834..146986,149628..149764,151186..151433,155052..155208,155948..155987))
/notes="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translation="MKLIMVLSLIAAAAEQNLVHGGPCDKTSHPYQALYTSGH
LLCGVLHPDWLWTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
HDQDMLRLARPAKLSLIQPLERDCSANTTSCHILGWGTADGFPDITQAYI
HLVRECEHAYPQGITQNMLCAGDERYKDCQCGSLGVLVCGDHLRGLVSWGNI
GSEKPGVYTVNCRYTNWIKTIQAK"
/complement(<165420..>170283)
/gene="KLK7"
/notes="synonym: SCCE"
/complement(join(<165420..165575,167672..167808,168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
/complement(join(165420..165575,167672..167808,168124..168371,169651..169798,170211..170283))
/gene="KLK7"
/notes="serine protease"
gene
/notes="serine protease"
Qy 1 GACACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACCGACCCATCTCTGC 60
Db 136548 GACACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACCGACCCATCTCTGC 136489
Qy 61 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTGTGACAGTGCACCTTGCCTAAG 120
Db 136488 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTGTGACAGTGCACCTTGCCTAAG 136429
Qy 121 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAGTGCAGAGTCTTACCCGA 180
Db 136428 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAGTGCAGAGTCTTACCCGA 136369
Qy 181 GACAGATAGATGACACCATGTTCTGCGCGTGCAGAACAGTAGACTCTCTGCCAGG 240
Db 136368 GACAGATAGATGACACCATGTTCTGCGCGGTGCAGAACAGTAGACTCTCTGCCAGG 136309
Qy 241 TGAGGACACCTCTCTTTATTACGAGATACACACTGAGTGCACACTCGGTAACTGGAGC 300
Db 136308 TGAGGACACCTCTCTTTATTACGAGATACACACTGAGTGCACACTCGGTAACTGGAGC 136249
Qy 301 GTTGCCAAATCTGAGAAATCCAGCAATTCAGAGAGTCCAGAGTCTTCTCAGAGA 360
Db 136248 GTTGCCAAATCTGAGAAATCCAGCAATTCAGAGAGTCCAGAGTCTTCTCAGAGA 136189
Qy 361 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGTCTGTTTATGTCATTTC 420
Db 136188 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGTCTGTTTATGTCATTTC 136129
Qy 421 CAGTTTTTATAGTACCCACATTTAAAAAGGTTGGGGCAGTGGCTCACAACCT 480
```

mRNA

CDS

mRNA

CDS

gene

mRNA

CDS

gene

mRNA

CDS

mRNA

CDS

gene

mRNA

CDS

gene

mRNA

CDS

gene

mRNA

CDS

```
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translation="MARSLLPLQILLLSLALETAGEAOGDKIIDGAPCARSGHPWQ
VALISGQLHCGGVLVNERWLTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
GSLYQTHVNDMLVKNLSQARLSMVKVRLPSRCEPPTCTVSGWGTTPSPDPTFP
SDLMCDVKLIISPDCTKVKYKLLLENSMLCAGIPDSKKNACNGDSGGPLVCRGTLOGL
VSMGTFCGPNPDGVTYQVCKFTKINDTWKKHR"
/complement(join(<183943..184098,185635..185768,187865..188127,188293..188452,188967..>189036))
/product="neuropsin"
/complement(join(183943..184098,185635..185768,187865..188127,188293..188452,188967..189036))
/notes="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translation="MGRPRPRAAKTWMFLLLGGAWAGHRAQBDKVLGGHECOPHSQ
PWOALPOGOQLLCGGVLVGNWLVTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
SIHPCVNSSDVEDHNDMLMLQLRDOASIGSKVKPISLADHCTQPCQKCTVSGWGT
TSRENPPDLNCAEVKIFQKCEDAYPQGITDGMVCAGSSKAGATCGDSSGGPLVC
DGAQGITSGSDPCGRSDRPGVYTNICRILDMIKIISRGK"
/complement(join(<190980..191129,191573..191709,194324..194589,197048..197204,197370..>197412))
/product="kallikrein-like 3"
/complement(join(190980..191129,191573..191709,194324..194589,197048..197204,197370..197412))
/notes="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translation="MKLGLCALLSLAHGWADTRAIGAEBCPNQWQAGLPHLT
RLFCGATLISDRMLLTAACHLKNQSVLGRNLPEPDTGORVPSHSFPHPLYN
LSANDHNDMLRLPRQARLSPAVQINLSQTCSPGMOQLISGWGAVSPKALFV
TLQCANISILENKLCHWAYPGHISDSMLCAGLWEGGRGSCQDSSGGPLVNCNTLAGVV
SGGAEPSCRPRPRAVYTSVCHYLDWIQIEMEN"
Query Match 100.0%; Score 1001; DB 9; Length 2300000;
Best Local Similarity 100.0%; Pred. No. 1.3e-227;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACCGACCCATCTCTGC 60
Db 136548 GACACCAAGAGCCCCAAGGTGAGTCCAGGTTCTTGTATACCGACCCATCTCTGC 136489
Qy 61 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTGTGACAGTGCACCTTGCCTAAG 120
Db 136488 CGCCTTCCATCTTCTCCACTTCTCATTTGTCTTGTGACAGTGCACCTTGCCTAAG 136429
Qy 121 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAGTGCAGAGTCTTACCCGA 180
Db 136428 TCCTCCAGTGTGAATATCAGCGTGTCTAGTCAGAAAAGTGCAGAGTCTTACCCGA 136369
Qy 181 GACAGATAGATGACACCATGTTCTGCGCGTGCAGAACAGTAGACTCTCTGCCAGG 240
Db 136368 GACAGATAGATGACACCATGTTCTGCGCGGTGCAGAACAGTAGACTCTCTGCCAGG 136309
Qy 241 TGAGGACACCTCTCTTTATTACGAGATACACACTGAGTGCACACTCGGTAACTGGAGC 300
Db 136308 TGAGGACACCTCTCTTTATTACGAGATACACACTGAGTGCACACTCGGTAACTGGAGC 136249
Qy 301 GTTGCCAAATCTGAGAAATCCAGCAATTCAGAGAGTCCAGAGTCTTCTCAGAGA 360
Db 136248 GTTGCCAAATCTGAGAAATCCAGCAATTCAGAGAGTCCAGAGTCTTCTCAGAGA 136189
Qy 361 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGTCTGTTTATGTCATTTC 420
Db 136188 GCTCATACCTTAGAGTAGTGTGTTTAGTGAATAATGCTGAGTCTGTTTATGTCATTTC 136129
Qy 421 CAGTTTTTATAGTACCCACATTTAAAAAGGTTGGGGCAGTGGCTCACAACCT 480
```







Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,  
Stantropop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,  
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 200792)  
Green, E.D.  
Direct Submission  
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 200792)  
Green, E.D.  
Direct Submission  
Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
On Nov 22, 2002 this sequence version replaced gi:22218452.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: nisc.zoonhgri.nih.gov  
----- Project Information  
Center project name: dhz  
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 196656 bases at least Q40  
Consensus quality: 197883 bases at least Q30  
Consensus quality: 198879 bases at least Q20  
Insert size: 215000; agarose-fp  
Insert size: 199692; sum-of-contigs  
Quality coverage: 9.03x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

1 28306: contig of 28306 bp in length  
\* 28307 28406: gap of unknown length  
\* 28407 37856: contig of 9450 bp in length  
\* 37857 37956: gap of unknown length  
\* 37957 73522: contig of 35566 bp in length  
\* 73523 73623: gap of unknown length  
\* 73623 83567: contig of 9945 bp in length  
\* 83568 83667: gap of unknown length  
\* 83668 88917: contig of 5150 bp in length  
\* 88918 88917: gap of unknown length  
\* 88918 125611: contig of 36694 bp in length  
\* 125612 159879: gap of 34168 bp in length  
\* 159880 159979: gap of unknown length  
\* 159980 174698: contig of 14719 bp in length

\* 174699 174798: gap of unknown length  
\* 174799 186382: contig of 11584 bp in length  
\* 186383 186482: gap of unknown length  
\* 186483 193344: contig of 6862 bp in length  
\* 193345 193444: gap of unknown length  
\* 193445 193663: contig of 5919 bp in length  
\* 193664 194663: gap of unknown length  
\* 194664 200792: contig of 1329 bp in length.  
FEATURES  
source  
location/Qualifiers  
1..200792  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="CH251-355A20"  
/clone\_lib="CH251"  
1..28306  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
misc\_feature  
28407..37856  
/note="assembly\_fragment"  
37957..73522  
/note="assembly\_fragment"  
73623..83567  
/note="assembly\_fragment"  
83668..88917  
/note="assembly\_fragment"  
88918..125611  
/note="assembly\_fragment"  
125712..159879  
/note="assembly\_fragment"  
159980..174698  
/note="assembly\_fragment"  
174799..186382  
/note="assembly\_fragment"  
186483..193344  
/note="assembly\_fragment"  
193445..199363  
/note="assembly\_fragment"  
199464..200792  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right

ORIGIN

Query Match 95.8%; Score 958.8; DB 2; Length 200792;  
Best Local Similarity 98.6%; Pred. No. 1.5e-217;  
Matches 988; Conservative 0; Mismatches 12; Indels 2; Gaps 2;  
Qy 1 GACACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGTATACCGACCATCTCTGC 60  
Db 73157 GACACCAAGAGCCCCCAAGGTGAGTGTCCAGGTTCTTTGTATATGACCATCTCTGC 73098  
Qy 61 CGCCTTCATCTTCTCCACTTCTCATTTGTTCTCTGTTGACAGTGCACATCTCCCTAAGG 120  
Db 73097 CGCCTTCATCTTCTCCACTTCTCATTTGTTCTCTGTTGACAGTGCCTTCCCTAAGG 73038  
Qy 121 TCCTCCAGTGTCTGAATATCAGCGTGTAAAGTCAGAAAAGGTGCGAGGATGCTTACCCGA 180  
Db 73037 TCCTCCAGTGTCTGAATATCAGCGTGTAAAGTCAGAAAAGGTGCGAGGATGCTTACCCGA 72978  
Qy 181 GACAGATAGATGACACCATGTTCTTGGCCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 240  
Db 72977 GACAGATAGATGACACCATGTTCTTGGCCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 72918  
Qy 241 TGAGGACACCTCTCTTTTATTCAGCAGATACACATGAGTGCACACTCGTAACTGAGGC 300  
Db 72917 TGAGGACACCTCTCTTTTATTCAGCAGATACACATGAGTGCACACTCGTAACTGAGGC 72858  
Qy 301 GTTCCCAAATTCAGAGATCCAGCAATTCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 360  
Db 72857 GTTCCCAAATTCAGAGATCCAGCAATTCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 72798





```

RESULT 10
AC133961/c
LOCUS       AC133961             177787 bp    DNA    linear    PRI 03-JAN-2003
DEFINITION  Homo sapiens BAC clone RP13-494C23 from 4, complete sequence.
ACCESSION   AC133961
VERSION     AC133961.3   GI:26190601
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1   (bases 1 to 177787)
Sulston,J.E. and Waterston,R.
TITLE       Toward a complete human genome sequence
JOURNAL     Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE     99063392
PUBMED      9847074
2   (bases 1 to 177787)
Isak,A., Bielicki,L., Creason,K. and Cotton,M.
TITLE       The sequence of Homo sapiens BAC clone RP13-494C23
JOURNAL     Unpublished (2001)
3   (bases 1 to 177787)
Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (20-SEP-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4   (bases 1 to 177787)
Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (05-NOV-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5   (bases 1 to 177787)
Waterston,R.H.
TITLE       Direct Submission
JOURNAL     Submitted (07-DEC-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
6   (bases 1 to 177787)
Waterston,R.
TITLE       Direct Submission
JOURNAL     Submitted (03-JAN-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Dec 7, 2002 this sequence version replaced gi:24580489.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
-----
Summary Statistics
-----
Center project name: H_FH0494C23
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01HG01165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.

Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1&2 or either MboI or DpnII for library segments 3&4. Size selected DNA was cloned into the pBACe3.6 vector between the EcoRI sites for library segments 1&2 or the BamHI sites for library segments 3&4. The ligation products were transformed into DH10B electrocompetent cells (BRL life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

#### NEIGHBORING SEQUENCE INFORMATION:

The sequence from base 67288 to base 67340 was derived from one plasmid subclone.

Polymorphisms have been identified between AC020706 and this sequence.

Data from AC020706 was used to finish this clone.

This sequence is not the entire insert of the clone. This clone is overlapped by AC020706 and AC093660.

FEATURES	source	Location/Qualifiers
	1..177787	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="4"
		/map="4"
		/clone="RP13-494C23"
		/clone_lib="RPCI-13"
	4..70	
repeat_region		/rpt_family="MIR"
repeat_region	703..833	
repeat_region		/rpt_family="Alu"
repeat_region	852..875	
repeat_region		/rpt_family="AT-rich"
repeat_region	911..1138	
repeat_region		/rpt_family="Alu"
repeat_region	1157..1191	
repeat_region		/rpt_family="(CAAAA)n"
repeat_region	1341..1408	
repeat_region		/rpt_family="GA-rich"
repeat_region	1432..1577	
repeat_region		/rpt_family="GA-rich"
repeat_region	2115..2458	
repeat_region		/rpt_family="L2"
repeat_region	2562..2603	
repeat_region		/rpt_family="MER1_type"
repeat_region	2773..2868	
repeat_region		/rpt_family="MER1_type"
repeat_region	2932..3240	
repeat_region		/rpt_family="Alu"
repeat_region	3264..3462	
repeat_region		/rpt_family="MaLR"
repeat_region	3463..3760	
repeat_region		/rpt_family="Alu"
repeat_region	3761..3934	
repeat_region		/rpt_family="MaLR"
repeat_region	5508..5796	
repeat_region		/rpt_family="Alu"
repeat_region	6977..7240	
repeat_region		/rpt_family="Alu"

repeat_region	7250..7270	/rpt_family="AT_rich"	491	CACCTTTGGGAGGCTGAGCGAGCAGATCACCTTTGGTTCAGGAGTTTGAGACTAGCTGGC	550
repeat_region	7524..7661	/rpt_family="MIR"	58323	CACCTTTGGGAGGCGGAGCGAGGTGGATCACCTGAGGTTCAGGAGTTTCGAGACCGAGCTGGC	58264
repeat_region	7891..8423	/rpt_family="ERVL"	551	CAACATGCGGAAACTCTGTCTTAAAAAATAACAAAAATTTAGCTTCGCTGGCATGGTGGCGG	610
repeat_region	9624..9825	/rpt_family="MIR"	58263	CAACATGCGGAAACCTGTCTCT-ACTGAAATACAAAATTTAGCTGGCATGTTGGCAC	58205
repeat_region	10105..10163	/rpt_family="L2"	611	GGGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACACAAGAAATCACTTTAAACCCAGGAGG	670
repeat_region	10414..10721	/rpt_family="Alu"	58204	ATGCGTGTGATCTCAGCTGCTCAGGAGGCTGAGCGAGGAGAAATCGTAGAACCCGAGGAGG	58145
repeat_region	11532..11671	/rpt_family="L1"	671	TGGAGGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT	729
repeat_region	12367..12389	/rpt_family="AT_rich"	58144	CGGAGGTTGCAATGAGCGGAGATTGCTTATGCTCAGCTCCAGCTGGGCGAGACGAGCGAGA	58085
repeat_region	13174..13474	/rpt_family="Alu"	730	CTTTTGTCTCAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAGTTAACTTTA	789
repeat_region	13507..13830	/rpt_family="Alu"	58084	CTCCATCTCAAAAAAAGAAAAAAGGTAAGAGAGCAGCGCATATGAATA	58025
repeat_region	13834..13969	/rpt_family="Alu"	790	ATAACCCCAATCTAT-----CCCAATACAAATCACTTTCAAAAGTGAATTAATA	836
repeat_region	14004..14059	/rpt_family="polypyrimidine"	58024	TTAATGTAATTTTAACTCAGTGTATCCAAATATTGTCATGTCAACATACAATTAATA	57965
repeat_region	16122..16152	/rpt_family="(T)n"	837	TAAACAAATTAATGATGAGTACTTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTGA	896
repeat_region	16824..17116	/rpt_family="Alu"	57964	TAAA--ATTTTAAATGAGATATTTTGAGCC--CTTTTTTTTATACAAAGGGTTTGA	57911
repeat_region	17122..17421	/rpt_family="Alu"	897	AAGTGAGTATATGTTATGCTGACGACACATCTCAATTTGAGCTAGCTACATTTTCAGGT	956
repeat_region	17422..18113	/rpt_family="Alu"	57910	ANTTCAATGATATTTTACACTTACAGGACATCTCAATTCAGACCGCCACATTAAAGT	57851
repeat_region	18681..18727	/rpt_family="L1"	957	GCTCAGTA 964	
repeat_region	20147..20162	/rpt_family="L2"	57850	GCTTAATA 57843	
repeat_region	20163..20386	/rpt_family="MIR"			
repeat_region	20387..20502	/rpt_family="MER1_type"			
repeat_region	20503..20614	/rpt_family="MIR"			
repeat_region	20617..20782	/rpt_family="(TA)n"			
repeat_region	21217..21523	/rpt_family="MIR"			
repeat_region	21726..21824	/rpt_family="L2"			
repeat_region	21947..22277	/rpt_family="CR1"			
repeat_region	24511..24634	/rpt_family="MER1_type"			
repeat_region	24922..25153	/rpt_family="L2"			
repeat_region	25154..25413	/rpt_family="MER1_type"			
repeat_region	25414..25462	/rpt_family="Alu"			
Query Match	27.7%;	Score 276.8;	DB 9;	Length 177787;	
Best Local Similarity	68.9%;	Pred. No. 3.7e-55;			
Matches 460;	Conservative 0;	Mismatches 187;	Indels 21;	Gaps 5;	
Qy	311	TCTGAGATCCAGCAATGCCAAGACAGTCTAGGACCCCTGTTCTTCACAGAGCTCATACCC	370		
Db	58503	TCTGGCTATTTTCTTTTCCCAAGAAAGTCTAGCTGACCTCAAAATTTAGAAATAGCACTATC	58444		
Qy	371	TAGAGTAGTGTGTTTAGTAAGAAATAGTGTGAGCTGCTTATGTTCATTTCCAGTTTCTTA	430		
Db	58443	CAGTTGAAATACAATGTGAACCAAAATTTCAAGCCACATATATCAATTTAAATTTCTCTA	58384		
Qy	431	GTAGCCATTAACAGGTAAGAAAGGCTGGGCGAGTGGCTCAGCTGCTCAGACCTGTAATCCAG	490		
Db	58383	GTAGCTATGTTAAAGAAAGTAAGAGAGCGCGGGGCGAGTGGCTCAGACCTGTAATCTTAG	58324		
<p>RESULT 11 HSDJ71H19 LOCUS DEFINITION Human DNA sequence from clone RPI-71H19 on chromosome 6 Contains part of the PRIM2A gene encoding DNA primase large subunit (p58), ESTs, STSs and GSSs, complete sequence. ACCESSION AL121958 VERSION AL121958.6 GI:8247268 KEYWORDS HTG; PRIM2A. SOURCE Homo sapiens (human) ORGANISM Homo sapiens REFERENCE 1 (bases 1 to 138411) AUTHORS Brown, J. TITLE Direct Submission JOURNAL Submitted (27-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk requests: clonerequest@sanger.ac.uk COMMENT On Jun 4, 2000 this sequence version replaced gi:8218079. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information</p>					



on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6>  
RPI-71H19 is from the library RPCI-1 constructed at the Roswell  
Park Cancer Institute by the group of Pieter de Jong. For further  
details see <http://bacpac.med.buffalo.edu/>  
VPCOR: PCYPAC2

This sequence is the entire insert of clone RPI-71H19 The true left  
end of clone RPI-216J23 is at 87497 in this sequence. The true  
right end of clone RPI-422B11 is at 29670 in this sequence.

## FEATURES

source	1..138411 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="6" /clone="RPI-71H19" /clone_lib="RPCI-1" complement(477..1095) /note="match: GSS: Em:AQ506907" complement(609..1046) /note="match: GSS: Em:AQ437147" complement(620..1089) /note="match: GSS: Em:AQ819330" complement(621..1083) /note="match: GSS: Em:AQ437147" complement(761..1000) /note="match: GSS: Em:B71825" complement(977..1090) /note="match: GSS: Em:AQ552044" 1514..1858 /note="match: GSS: Em:AQ176950" 1545..1851 /note="match: GSS: Em:AQ176950" 1852..2161 /note="AluX repeat: matches 1..309 of consensus" 2510..8661 /note="L1PA4 repeat: matches 13..6142 of consensus" 8448..8976 /note="match: GSS: Em:AQ347835" 8662..8847 /note="match: GSS: Em:AQ347835" 8848..9156 /note="MLTII repeat: matches 40..410 of consensus" 9172..9481 /note="AluSg repeat: matches 5..313 of consensus" 9692..9804 /note="MLTII repeat: matches 267..371 of consensus" 9824..10043 /note="MIR repeat: matches 18..251 of consensus" 10338..10689 /note="MLT2PA repeat: matches 9..356 of consensus" complement(10977..11390) /note="match: STS: Em:G42414" complement(11825..12306) /note="match: GSS: Em:AQ141424" 12402..12825 /note="match: GSS: Em:AQ266412" 12647..12722 /note="MIR repeat: matches 1..79 of consensus" 13978..14287 /note="AluDb repeat: matches 1..310 of consensus" 14664..15035 /note="L2 repeat: matches 2294..2673 of consensus" 15036..15342 /note="AluSg repeat: matches 1..308 of consensus" 15343..15421 /note="L2 repeat: matches 2673..2750 of consensus" 15485..15534 /note="MIR repeat: matches 213..262 of consensus"
repeat_region	15581..15703 /note="MIR repeat: matches 110..234 of consensus" 16310..16535 /note="MIR repeat: matches 21..231 of consensus" 16597..17598 /note="L1PA7 repeat: matches 5135..6145 of consensus" 17617..17827 /note="AluSg repeat: matches 83..309 of consensus" 17828..19349 /note="L1PA7 repeat: matches 3624..5146 of consensus" complement(19723..20194) /note="match: GSS: Em:A2016849" 20519..20665 /note="match: GSS: Em:AQ371979" 20807..21136 /note="match: GSS: Em:AQ508843" 20807..21096 /note="match: GSS: Em:AQ508843" 20818..21272 /note="match: GSS: Em:AQ209413" 20818..21096 /note="match: GSS: Em:AQ209413" 20831..21139 /note="match: GSS: Em:AQ514718" 20902..21139 /note="match: GSS: Em:AQ093115" 21097..21176 /note="40 copies 2 mer ct 76% conserved" 21524..22065 /note="match: GSS: Em:AQ530328" 22856..23094 /note="AluSg repeat: matches 1..239 of consensus" 23102..23135 /note="17 copies 2 mer aa 88% conserved" 24117..24204 /note="MER5B repeat: matches 1..90 of consensus" 24883..25185 /note="AluY repeat: matches 1..300 of consensus" 25364..25449 /note="MSTD repeat: matches 1..96 of consensus" 25932..26300 /note="L1MB8 repeat: matches 5779..6162 of consensus" 26483..26600 /note="L2 repeat: matches 2583..2703 of consensus" 27463..27633 /note="AluSg repeat: matches 137..308 of consensus" complement(27586..28035) /note="match: GSS: Em:AQ753792" complement(27650..27736) /note="match: GSS: Em:AQ753792" 27737..28017 /note="AluJo repeat: matches 1..276 of consensus" 28766..28860 /note="MIR repeat: matches 85..188 of consensus" 29043..29225 /note="FRAM repeat: matches 0..174 of consensus" 29387..29497 /note="MIR repeat: matches 29..143 of consensus" 29585..29879 /note="AluSg repeat: matches 1..296 of consensus" 31001..31343 /note="L1MD2 repeat: matches 5957..6332 of consensus" 31344..31643 /note="AluSc repeat: matches 1..301 of consensus" 31644..31712 /note="L1MD2 repeat: matches 5890..5957 of consensus" 31714..31863 /note="L1MD3 repeat: matches 7589..7735 of consensus" 31874..32598 /note="L1MD2 repeat: matches 5386..6122 of consensus" 33641..33934 /note="AluX repeat: matches 3..298 of consensus" 33933..33941









## AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albrooks, S.L., Amaraturunge, H.C., Are, J.R., Banks, T., Barbarella, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burkett, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulsged, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguri, M., Okwunodu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Perez, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scher, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telirod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

## TITLE

Direct Submission

## REFERENCE

2 (bases 1 to 177876)

## AUTHORS

Worley, K.C.

## TITLE

Direct Submission

Submitted (07-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 25, 2001 this sequence version replaced gi:12831249.

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: HAMH  
Center clone name: RP11-533F2  
----- Sequencing Statistics  
Sequencing vector: M13; L08821  
Chemistry: Dye-terminator Big Dye; 99% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 169802 bases at least Q40  
Consensus quality: 180431 bases at least Q30  
Consensus quality: 185136 bases at least Q20  
Estimated insert size: 177736; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 42878: contig of 42878 bp in length  
\* 42879 42978: gap of unknown length  
\* 42979 62018: contig of 19040 bp in length  
\* 62019 62118: gap of unknown length  
\* 62119 80603: contig of 18485 bp in length  
\* 80604 80703: gap of unknown length  
\* 80704 82888: contig of 17585 bp in length  
\* 82889 93388: gap of unknown length  
\* 93389 113934: contig of 15546 bp in length  
\* 113935 114034: gap of unknown length  
\* 114035 125339: contig of 11305 bp in length  
\* 125340 125439: gap of unknown length  
\* 125440 138329: contig of 12890 bp in length  
\* 138330 138429: gap of unknown length  
\* 138430 145577: contig of 7148 bp in length  
\* 145578 145677: gap of unknown length  
\* 145679 150952: contig of 5275 bp in length  
\* 150953 151052: gap of unknown length  
\* 151053 155923: contig of 4871 bp in length  
\* 155924 156024: gap of unknown length  
\* 156025 159748: contig of 3725 bp in length  
\* 159749 159848: gap of unknown length  
\* 159849 163146: contig of 3298 bp in length  
\* 163147 163246: gap of unknown length  
\* 163247 166618: contig of 3372 bp in length  
\* 166619 166718: gap of unknown length  
\* 166719 170864: contig of 4146 bp in length  
\* 170865 170964: gap of unknown length  
\* 170965 173099: contig of 2135 bp in length  
\* 173100 173199: gap of unknown length  
\* 173200 175517: contig of 2318 bp in length  
\* 175518 175617: gap of unknown length  
\* 175618 177876: contig of 2259 bp in length.

## FEATURES

## source

1. 177876  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-533F2"

## ORIGIN

Query Match	26.08;	Score	260;	DB 2;	Length	177876;
Best Local Similarity	67.74;	Pred. No.	3.7e-51;			
Matches	447;	Conservative	0;	Mismatches	190;	Indels 23; Gaps 5;

  

Qy	344	ACCCCTGTTCTCAGAGCTCATACCTAGAGTAGTGGTGTGTTAGTAGAATAATGCTGA	403
Db	22489	ACATCTAACATTACTAATCTCTAATCCAGGCTGTCCTCGAATGTGAGCCACATGATC	22548
Qy	404	GCTGCTTATGTCATTTCCAGTTTTTTAGTAGCCACATT-AAAACAGGTAAAGGCTGG	462
Db	22549	ACTACATAGGCAATTTTAAATTTATCTTGTAGGCACATTAAAAATAATAGTAGGCTGG	22608
Qy	463	CGCGAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCTGAGGAGGAGATCACCT	522
Db	22609	GTGCAGTGGCTCACACATGTAATCCAGCACTTTTCAAGAGCTAAGGTGGGAGATCACTT	22668
Qy	523	TTGTCAGGAGTTTGAGACATAGCTGCGCCCAACATGCGGAACCTGCTCTTAAAAAAA	582
Db	22669	GAGTCTAGGAGTTGAGAGCAGCTGCTGCAACATGGTGAACCCCGTCTCT-CTAAAAA	22727
Qy	583	TACAAAAATTAGCTGCGCTGCGGGGGCGCTGTAATCTCAGCTGCTCAGGAGGCGGA	642
Db	22728	CACAAAAATTAGCGGGTGTGGTGGCACACGCTGTAGTCTCAGCTACTCAGGAGGCTGA	22787
Qy	643	GACAAAGATCACTTAAACCCAGGAGGTTCGAGGCTCAGATCGATCGGCCACT	702





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 352.271 Seconds

(without alignments)  
16821.327 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_6000\_7000

Perfect score: 1001

Sequence: 1 gacaacaaagagccccaag.....gttactgtatggatggcac 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04.\*

1: geneseqn1980s.\*

2: geneseqn1990s.\*

3: geneseqn2000s.\*

4: geneseqn2001as.\*

5: geneseqn2001bs.\*

6: geneseqn2002as.\*

7: geneseqn2002bs.\*

8: geneseqn2003as.\*

9: geneseqn2003bs.\*

10: geneseqn2003cs.\*

11: geneseqn2003ds.\*

12: geneseqn2004as.\*

13: geneseqn2004bs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	3	AA95905
2	1001	100.0	11570	12	ADK52482
3	1001	100.0	11570	13	ADR72623
4	1001	100.0	11570	13	ADR72875
5	252.6	25.2	52216	4	AAH28355
6	252.6	25.2	52216	6	ABL50307
7	248.8	24.9	56098	11	ACN44922
8	240.2	24.0	55827	8	ACA60949
9	240.2	24.0	55827	10	ABX13671
10	240.2	24.0	58337	13	ADS36454
11	240.2	24.0	64423	13	ADS36462
12	235.6	23.5	27189	11	ACN45066
13	227.4	22.7	2953	4	AAK68359
14	227.4	22.7	2953	4	AAK68360
15	227.4	22.7	3377	12	ADQ64563
16	227.4	22.7	13744	4	AAK82406
17	227.4	22.7	13744	4	AAK68361
18	227.4	22.7	23071	4	AA826699
19	227.4	22.7	23071	8	AA98598
20	227.4	22.7	23071	8	ABX74048

21	227.4	22.7	23071	8	ADA44323	Ada44323 Human sec
22	227.4	22.7	23071	10	ADC20712	Adc20712 Human sec
23	227.4	22.7	23071	10	ADF10881	Adf10881 Human sec
24	226.8	22.7	12275	4	AAK85734	Aak85734 Human imm
25	226.2	22.6	133955	11	ACN45170	Acn45170 Human gen
26	225.4	22.5	325791	4	AA843104	Aa843104 Human Oes
27	225.2	22.5	35641	6	ABL64428	Ab164428 Stomach c
28	225.2	22.5	35641	6	ABN95727	Abn95727 Gene #222
29	224.8	22.5	629	12	ADM99987	Adm99987 Human STA
30	224.8	22.5	629	12	ADM99988	Adm99988 Human STA
31	224.6	22.4	110000	11	ACN43998_3	Continuation (4 of
32	223.8	22.4	15765	5	ABA18276	Abal18276 Human ner
33	223.8	22.4	15765	5	ABA19009	Abal19009 Human ner
34	223.8	22.4	15765	5	ABA17911	Abal17911 Human ner
35	223.8	22.4	15765	5	ABA18250	Abal18250 Human ner
36	223.6	22.3	12267	4	AAK85733	Aak85733 Human imm
37	222	22.2	72409	13	ABD33061	Abd33061 Human can
38	221.6	22.1	2591	4	AAH16224	Aah16224 Human cDN
39	221.6	22.1	2591	12	ADL22885	Adl22885 Human MP2
40	221.6	22.1	2591	13	ADR14082	Adr14082 Human NF-
41	220.8	22.1	96587	9	ADA02984	Ada02984 Human MAP
42	220.8	22.1	96587	10	ADB72722	Adb72722 Human MAP
43	220.8	22.1	96587	10	ADC85464	Adc85464 Human Map
44	220.8	22.1	96587	12	ADM74579	Adm74579 Human car
45	220.6	22.0	203264	11	ACN44506	Acn44506 Human gen

#### ALIGNMENTS

#### RESULT 1

AAA95905

ID AAA95905 standard; DNA; 11570 BP.

XX AAA95905;

DT 02-FEB-2001 (first entry)

XX Human KLK-L2 gene.

DE Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;

XX kallikrein-like protein; serine protease; cytostatic; cancer;

XX prostrate cancer; ds.

XX Homo sapiens.

OS WO200053776-A2.

PN 14-SEP-2000.

PD 09-MAR-2000; 2000WO-CA000258.

XX 11-MAR-1999; 99US-0124260P.

XX 01-APR-1999; 99US-0127386P.

XX 21-JUL-1999; 99US-0144919P.

XX (MOUN ) MOUNT SINAI HOSPITAL.

XX Yousef GM, Diamandis EP;

XX WPI; 2000-587440/55.

XX P-ESDB; AAB21296.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L

XX protein mediated disorders, especially cancer.

XX Claim 1; Page 143-149; 184pp; English.

XX The present sequence is the coding sequence of the human KLK-L2 gene,

XX which encodes a kallikrein-like protein. Kallikreins and kallikrein-like

XX proteins are a subgroup of the serine protease enzyme family. They

XX catalyse the selective cleavage of specific polypeptide precursors to

XX release peptides with potent biological activity. Nucleic acids encoding



Db 6120 TCCTCCAGTCTTGAATATCAGCGTGTAAAGTCAGAAAAAGGTGCGAGGATGCTTACCCGA 6179  
QY 181 GACAGATAGTACACACCATGTTCTCGCGCGGTGACAAAGCAGGTAGACATCTCTCCAGG 240  
Db 6180 GACAGATAGTACACACCATGTTCTCGCGCGGTGACAAAGCAGGTAGACATCTCTCCAGG 6239  
QY 241 TGAGGACACCTCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC 300  
Db 6240 TGAGGACACCTCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC 6299  
QY 301 GTTGCCAAATCTGAGATCCAGCAATTTGCCAAGACAGTTCAGGACCCCTGTTTCTCACAGA 360  
Db 6300 GTTGCCAAATCTGAGATCCAGCAATTTGCCAAGACAGTTCAGGACCCCTGTTTCTCACAGA 6359  
QY 361 GCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGAGCTGCTTATGTCATTTTC 420  
Db 6360 GCTCATACCTAGAGTAGTGGTGTAGTAGAATAATGCTGAGCTGCTTATGTCATTTTC 6419  
QY 421 CAGTTTTTTAGTAGCCACATTAACACAGGTAAAAAGGCTGGCGCAGTGGCTCACACT 480  
Db 6420 CAGTTTTTTAGTAGCCACATTAACACAGGTAAAAAGGCTGGCGCAGTGGCTCACACT 6479  
QY 481 GTAATCCAGCAGCTTTGGAGGCTCAGCAGCAGATCACCTTTGGTCAGGAGTTTGAGA 540  
Db 6480 GTAATCCAGCAGCTTTGGAGGCTCAGCAGCAGATCACCTTTGGTCAGGAGTTTGAGA 6539  
QY 541 CTAGCTGGCCAAACATGCGGAACTCTGTCTCTAAAAAATACAAAAAATAGCTGGC 600  
Db 6540 CTAGCTGGCCAAACATGCGGAACTCTGTCTCTAAAAAATACAAAAAATAGCTGGC 6599  
QY 601 ATGTGGCGGCGCTGTAACTCTCAGCTGCTCAGGAGCCGAGACACAGAAATCACTTAA 660  
Db 6600 ATGTGGCGGCGCTGTAACTCTCAGCTGCTCAGGAGCCGAGACACAGAAATCACTTAA 6659  
QY 661 ACCCAGGAGTGGAGGTTGCAGTGAGTGCAGTGCCTCACTCCCAACCTGGGAGAC 720  
Db 6660 ACCCAGGAGTGGAGGTTGCAGTGAGTGCAGTGCCTCACTCCCAACCTGGGAGAC 6719  
QY 721 AGAGTGACATTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAAG 780  
Db 6720 AGAGTGACATTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAAG 6779  
QY 781 TTAACCTTAATAACCAATGTATCCCAATCAATCAATCAATCAATCAATCAATCAATCAAT 840  
Db 6780 TTAACCTTAATAACCAATGTATCCCAATCAATCAATCAATCAATCAATCAATCAATCAAT 6839  
QY 841 ACAATTAATGATGAGATCTTTTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 900  
Db 6840 ACAATTAATGATGAGATCTTTTACATTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 6899  
QY 901 GAGTATATGTTATGCTGACAGCAGATCTCAATTTGGAGTACATTTTCAATTTCAATTTCAATTT 960  
Db 6900 GAGTATATGTTATGCTGACAGCAGATCTCAATTTGGAGTACATTTTCAATTTCAATTTCAATTT 6959  
QY 961 AGTAGCCACATGTGCTAGCAGTACTGTATTGGATGCCAC 1001  
Db 6960 AGTAGCCACATGTGCTAGCAGTACTGTATTGGATGCCAC 7000

RESULT 3  
ADR72623  
ID ADR72623 standard; DNA; 11570 BP.  
XX  
AC ADR72623;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.  
XX  
KW kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;  
KW cancer metastasis; chemotherapy; human; serine protease;  
KW chromosome 19q13.4; KLK5; ds; gene.  
XX

Homo sapiens.

OS  
XX  
FH Key  
FT CDS  
FT  
FT  
FT product= "Human renal cell carcinoma-related kallikrein  
5 (hK5) protein"  
2221..2293  
FT /tag= a  
FT /number= 1  
FT intron  
FT /tag= c  
FT /number= 1  
FT exon  
FT /tag= d  
FT /number= 2  
FT intron  
FT /tag= e  
FT /number= 2  
FT exon  
FT /tag= f  
FT /number= 3  
FT intron  
FT /tag= g  
FT /number= 3  
FT exon  
FT /tag= h  
FT /number= 4  
FT exon  
FT /tag= i  
FT /number= 4  
FT exon  
FT /tag= j  
FT /number= 5  
WO2004077060-A2.  
10-SEP-2004.  
26-FEB-2004; 2004WO-CA000280.  
27-FEB-2003; 2003US-0451382P.  
(MOUN ) MOUNT SINAI HOSPITAL.  
Diamandis EP, Petraki CD;  
WPI; 2004-662077/64.  
P-PSDB; ADR72621.

Detecting kallikrein polypeptides/polynucleotides associated with renal cell carcinoma in a patient, for diagnosing/treating the disease, comprises detecting /identifying kallikrein polypeptides/polynucleotides in a sample.

Example 1; SEQ ID NO 3; 53pp; English.

The invention relates to a novel method for detecting kallikrein polypeptides, or the polynucleotides encoding them, associated with renal cell carcinoma. The method comprises obtaining a sample from a patient and detecting kallikrein polypeptides, or their encoding polynucleotides, where the kallikrein polypeptides are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The detected amounts of the kallikrein polypeptides are compared to standard amounts. The molecules of the invention demonstrate cytostatic activity whilst the methods and kit may be useful for detecting, characterising, preventing and treating renal cell carcinoma. Furthermore, the methods may be useful for evaluating the probability of the presence of malignant or pre-malignant cells and for detecting and quantitating tumour growth and cancer metastasis. Finally, the methods may be utilised to confirm the absence or removal of all tumour tissue following surgery, cancer chemotherapy and/or radiation therapy and to monitor cancer chemotherapy and tumour reappearance. The current sequence is that of the human

CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the  
CC invention which encodes a secreted serine protease and is located at  
CC chromosome 19q13.4.  
XX  
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1001; DB 13; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 1e-205;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACACCAAGAGCCCCCAAGGTGAGTCCAGGTTCTTCTGTATACCGACCCATCTCTGC 60  
Db 6000 GACACCAAGAGCCCCCAAGGTGAGTCCAGGTTCTTCTGTATACCGACCCATCTCTGC 6059  
Qy 61 CGCCTTCATCTTCTCCACTTCTCATTTGTTCTGTTTGCAGTGCACATTCCTTAAGG 120  
Db 6060 CGCCTTCATCTTCTCCACTTCTCATTTGTTCTGTTTGCAGTGCACATTCCTTAAGG 6119  
Qy 121 TCCTCCAGTGTGGAATATCAGCTGCTTAAGTCTAGAAAAGGTGCGAGATGCTTTACCCGA 180  
Db 6120 TCCTCCAGTGTGGAATATCAGCTGCTTAAGTCTAGAAAAGGTGCGAGATGCTTTACCCGA 6179  
Qy 181 GACAGATAGATGACACCAATGTTCTGCGCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 240  
Db 6180 GACAGATAGATGACACCAATGTTCTGCGCGGTGACAAAGCAGGTAGAGATCTCTGCCAGG 6239  
Qy 241 TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCTGCTGTTTAAACATGAGC 300  
Db 6240 TGAGGACACCTCTCTTTATTCAGCAGATACACACTGAGTGCCTGCTGTTTAAACATGAGC 6299  
Qy 301 GTTGCCAAATTCAGAAATCCAGAAATGCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 360  
Db 6300 GTTGCCAAATTCAGAAATCCAGAAATGCGCAAGACAGTTCAGGACCCCTGTTCTCACAGA 6359  
Qy 361 GCTCATACCTTAGAGTAGTGTGTTTGTAGTAAATAATAGCTGAGTGTGTTATGTCATTTTC 420  
Db 6360 GCTCATACCTTAGAGTAGTGTGTTTGTAGTAAATAATAGCTGAGTGTGTTATGTCATTTTC 6419  
Qy 421 CAGTTTTTTAGTAGCCACATTAACAGGTAAAGGCTGGCGCAGTGGCTCACACCT 480  
Db 6420 CAGTTTTTTAGTAGCCACATTAACAGGTAAAGGCTGGCGCAGTGGCTCACACCT 6479  
Qy 481 GTAATCCAGACATTTGGGAGGCTGAGCAGGCGAGATCACCTTTGGTCAGGAGTTTGAGA 540  
Db 6480 GTAATCCAGACATTTGGGAGGCTGAGCAGGCGAGATCACCTTTGGTCAGGAGTTTGAGA 6539  
Qy 541 CTAGCCTGGCCAAACATGGCGAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCTGGC 600  
Db 6540 CTAGCCTGGCCAAACATGGCGAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCTGGC 6599  
Qy 601 ATGTTGGCGGGCGCTGTAAATCTCAGCTGCTCAGGAGCCGAGACACAGAAATCACTTAA 660  
Db 6600 ATGTTGGCGGGCGCTGTAAATCTCAGCTGCTCAGGAGCCGAGACACAGAAATCACTTAA 6659  
Qy 661 ACCCAGGAGTGGAGTTGCGAGTGCAGTGCAGTGCCTCACTCCCAACTGGGAGAC 720  
Db 6660 ACCCAGGAGTGGAGTTGCGAGTGCAGTGCAGTGCCTCACTCCCAACTGGGAGAC 6719  
Qy 721 AGAGTGCACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAGAAAAACAGGTGGAAG 780  
Db 6720 AGAGTGCACATTTTGTCTCAAAAAAGAAAAAACAAGTAAAGAAAAACAGGTGGAAG 6779  
Qy 781 TTAATTTAATAACCAATGATATCCCAATACAAATCAATTTCAAGGTGTAATTAATAAA 840  
Db 6780 TTAATTTAATAACCAATGATATCCCAATACAAATCAATTTCAAGGTGTAATTAATAAA 6839  
Qy 841 ACAATATGATGAGATCTTACATCTTTCTTTTCTTATTAAGTCTTTGGAAGT 900  
Db 6840 ACAATATGATGAGATCTTACATCTTTCTTTTCTTATTAAGTCTTTGGAAGT 6899  
Qy 901 GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC 960  
Db 6900 GAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGCTC 6959

Qy 961 AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC 1001  
Db 6960 AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC 7000

# RESULT 4

ADR72875  
ID ADR72875 standard; DNA; 11570 BP.

XX ADR72875;

XX XX 02-DEC-2004 (first entry)

XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.

XX kallikrein 5; tumour marker; ovarian cancer;

KW epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;

XX KLK5; ds; gene.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 2221..11247

FT FT /\*tag= b

FT FT /product= "Human ovarian cancer-related tumour marker

FT FT kallikrein 5 (hk5) protein"

FT FT 2221..2293

FT FT /\*tag= a

FT FT /number= 1

FT FT /\*tag= c

FT FT /number= 1

FT FT 4762..5023

FT FT /\*tag= d

FT FT /number= 2

FT FT 5024..5762

FT FT /\*tag= e

FT FT /number= 2

FT FT 5763..6019

FT FT /\*tag= f

FT FT /number= 3

FT FT 6020..6104

FT FT /\*tag= g

FT FT /number= 3

FT FT 6105..6238

FT FT /\*tag= h

FT FT /number= 4

FT FT 6239..11091

FT FT /\*tag= i

FT FT /number= 4

FT FT 11092..11247

FT FT /\*tag= j

FT FT /number= 5

XX WO2004075713-A2.

XX PD 10-SEP-2004.

XX 26-FEB-2004; 2004WO-CA000281.

XX 26-FEB-2003; 2003US-0450406P.

XX (MOUN ) MOUNT SINAI HOSPITAL.

XX Diamandis EP;

XX WPI; 2004-661815/64.

XX P-ESDB; ADR72873.

XX Kallikrein markers detection method for detecting ovarian cancer in patient, involves detecting kallikrein markers and CA125 in sample obtained from patient, and comparing detected amounts with standard

amounts.

XX PS Example 2; SEQ ID NO 5; 102pp; English.

XX CC The invention relates to a novel method for detecting a plurality of kallikrein markers associated with ovarian cancer. The method comprises obtaining a sample from a patient and detecting in the sample a plurality of kallikrein markers, and optionally carbohydrate antigen CA125, wherein the kallikrein markers are selected from the group consisting of CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and CC kallikrein 11. The detected amounts of the kallikrein markers are CC compared with standard amounts. The method of the invention may be useful CC for detecting kallikrein markers associated with ovarian cancer in a CC patient and thus for detecting ovarian cancer, particularly epithelial CC ovarian carcinoma. The current sequence is that of the human ovarian CC cancer-related tumour marker kallikrein 5 (KUK5) DNA 2 of the invention CC which encodes a secreted serine protease and is located at chromosome CC 19q13.4.

XX SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1001;	DB 13;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 1e-205;		
Matches 1001; Conservative	-0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GACAAACAGAGCCCCCAAGTGAGTGTCCAGAGTTCTTCTTGATACGACCCCATCTCTGC	60	
Db	6000	GACAAACAGAGCCCCCAAGTGAGTGTCCAGAGTTCTTCTTGATACGACCCCATCTCTGC	6059	
Qy	61	CGCCTTCCATCTTTTCCCATCTTCATGTGTTTCTCTGTTTGACAGTGCATCTCCCTAAGG	120	
Db	6060	CGCCTTCCATCTTTTCCCATCTTCATGTGTTTCTCTGTTTGACAGTGCATCTCCCTAAGG	6119	
Qy	121	TCCTCCAGTGTGGAATATCAGCGTGTCTAGTCAAGAAAGTGCGGAGGATGCTTACCCGA	180	
Db	6120	TCCTCCAGTGTGGAATATCAGCGTGTCTAGTCAAGAAAGTGCGGAGGATGCTTACCCGA	6179	
Qy	181	GACAGATAGATGACACCATGTCTTCGCCCGTGACAAAGCAGGTAGAGATCTCTGCCAGG	240	
Db	6180	GACAGATAGATGACACCATGTCTTCGCCCGTGACAAAGCAGGTAGAGATCTCTGCCAGG	6239	
Qy	241	TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	300	
Db	6240	TGAGGACACCTCTCTTTTATTTCAGCAGATACACACTGAGTGCCAACTCGGTAAACATGGAGC	6299	
Qy	301	GTTGCCAAATTCGAGAAATCCAGCAATTCGACAGACAGTCAAGACCCCTGTTCTCAGAGA	360	
Db	6300	GTTGCCAAATTCGAGAAATCCAGCAATTCGACAGACAGTCAAGACCCCTGTTCTCAGAGA	6359	
Qy	361	GCTCATACCTTCAGAGTAGTGGTGTATTAGTAGAATAATGCTGAGCTGCTTATGTCTATTTC	420	
Db	6360	GCTCATACCTTCAGAGTAGTGGTGTATTAGTAGAATAATGCTGAGCTGCTTATGTCTATTTC	6419	
Qy	421	CAGTTTTTTTAGTAGCCACATTTAAACAGGTAAACAAAGCGTGGCGCAGTGGCTCACACCT	480	
Db	6420	CAGTTTTTTTAGTAGCCACATTTAAACAGGTAAACAAAGCGTGGCGCAGTGGCTCACACCT	6479	
Qy	481	GTAATCCAGGACHTTGGGAGGCTGAGCGCAGGCGAGATCACTTTTGGTTCAGGAGTTTGA	540	
Db	6480	GTAATCCAGGACHTTGGGAGGCTGAGCGCAGGCGAGATCACTTTTGGTTCAGGAGTTTGA	6539	
Qy	541	CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATAACAAAAATAGCCTGGC	600	
Db	6540	CTAGCCTGGCCAAACATGGCGAAACTCTGTCTCTAAAAAAAATAACAAAAATAGCCTGGC	6599	
Qy	601	ATGTTGGCGGGCGCTGTGAATCTCAGCTGTCTCAGGAGCGCGAGACACAAGAAATCACTTAA	660	
Db	6500	ATGTTGGCGGGCGCTGTGAATCTCAGCTGTCTCAGGAGCGCGAGACACAAGAAATCACTTAA	6659	
Qy	661	ACCAGAGAGTGGAGTTTGAGTGAAGTCTGTCAGTGTCTCAGGAGCGCGAGACACAACCTGGGAGAC	720	
Db	6660	ACCAGAGAGTGGAGTTTGAGTGAAGTCTGTCAGTGTCTCAGGAGCGCGAGACACAACCTGGGAGAC	6719	

Qy	721	AGAGTGACACCTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAAG	780
Db	6720	AGAGTGACACCTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGAAG	6779
Qy	781	TTAACTTTTAATAACCCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATTAATATAA	840
Db	6780	TTAACTTTTAATAACCCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATTAATATAA	6839
Qy	841	ACAATTATGAATGAGATACCTTTACATCTCTTTCTTGTCTTTTCATATTAAAGTCTTTGAAAGT	900
Db	6840	ACAATTATGAATGAGATACCTTTACATCTCTTTCTTGTCTTTTCATATTAAAGTCTTTGAAAGT	6899
Qy	901	GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGTCTC	960
Db	6900	GAGTATATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTCAAGTGTCTC	6959
Qy	961	AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC	1001
Db	6960	AGTAGCCACATGTGGCTAGCAGTTACTGTATTGGATGGCAC	7000
RESULT 5			
AAH28355			
ID	AAH28355	standard; DNA; 52216 BP.	
XX	AAH28355;		
XX	05-SEP-2001	(first entry)	
XX	Nucleotide sequence of the human musashi promoter.		
XX	Musashi promoter; multipotential neural progenitor cell;		
KW	neural stem cell; central nervous system; ss.		
XX	Homo sapiens.		
OS	WO200146384-A2.		
XX	28-JUN-2001.		
XX	22-DEC-2000; 2000WO-US035395.		
XX	23-DEC-1999; 99US-0173003P.		
PR	(CORR ) CORNELL RES FOUND INC.		
PA	(NISC-) JAPAN SCI & TECHNOLOGY CORP.		
PA	Goldman SA, Okano H;		
FI	WPI; 2001-418053/44.		
XX	Separating multipotential neural progenitor cells from a mixed population		
XX	of cells, involves introducing nucleic acid molecule encoding fluorescent		
PT	protein under promoter control, and separating fluorescent cells.		
PT	Claim 12; Fig 11A-JJ; 87pp; English.		
PS	The present sequence represents the human musashi promoter, which is used		
XX	in the method of the invention. The specification describes a method for		
CC	separating multipotential neural progenitor cells from a mixed population		
CC	of cell types. The method comprises introducing a nucleic acid encoding a		
CC	fluorescent protein under control of a promoter which selectively		
CC	functions in the progenitor cells, into all cell types of the mixed		
CC	population, allowing only the progenitor cells to express the fluorescent		
CC	protein, and identifying and separating cells that are fluorescent, where		
CC	the separated cells are progenitor cells. The method is useful for		
CC	isolation and purification of multipotential neural progenitor cells,		
CC	especially neural stem cells from adult brain. The isolated cells are		
CC	used in both basic analyses of precursor and stem cell growth control, as		
CC	well as in more applied studies of their transplantability and		
CC	engraftment characteristics. The cells are useful in support of the		
CC	structural repair of the damaged central nervous system, such as in the		
CC	traumatized brain, or the contoured,traumatized or transected spinal		



## RESULT 7

ACN44922/c  
ID ACN44922 standard; DNA; 56098 BP.

XX AC

ACN44922;

XX AC

DT 18-NOV-2004 (first entry)

XX DT

DE Human genomic sequence HCG1779020.

XX DE

KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX KW

OS Homo sapiens.

XX OS

FN WO2003073826-A2.

XX FN

PD 12-SEP-2003.

XX PD

PF 28-FEB-2003; 2003WO-US006235.

XX PF

PR 01-MAR-2002; 2002US-00087192.

XX PR

PA (SAGR-) SAGRES DISCOVERY.

XX PA

PI Morris DW;

XX PI

DR WPI; 2003-328604/31.

XX DR

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX PT

PS comprises a nucleotide sequence.

XX PS

Claim 1; SEQ ID NO 1612; opp; English.

XX Claim

The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

XX US2002182586A1, for which no sequence data was published

SQ Sequence 56098 BP; 17842 A; 10579 C; 10322 G; 16942 T; 0 U; 413 Other;

XX SQ

Query Match 24.9%; Score 248.8; DB 11; Length 56098;  
Best Local Similarity 71.0%; Pred. No. 1.2e-43;  
Matches 402; Conservative 0; Mismatches 152; Indels 12; Gaps 5;

QY 440 TTAACAAGGTAAAGGCTGGCGCAGTGGCTCACACCTGTATATCCAGCATTGGG 499  
Db TTTAAAACTGATGACAGCCAGGCGTGGTGGCTCACACCTGTATATCCAGCATTGGG 26621

QY 500 AGGCTGAGGAGGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCACATGCC 559  
Db AGGCTGAGGTTGGGAGATCACCTTTGGTCAGGAGTTTGAGAAATTTGAAACAGCCCTGGCCACATGGT 26561

QY 560 GAAACTCTGTCTCTAAAAAATAACAAAAATTTAGCCTGGCATGTGGCGGCGCTCTA 619  
Db GAAACCCCATCTCT-CTAAATAATAGAAAAGTTAGCCGATGTGGTGCAACCGTTGTA 26502

QY 620 ATCTCAGCTGCTCAGGAGCCGAGACACAAGAAATCACTTAAACCCAGGAGTGGAGTTG 679  
Db GTCCAGACGCTCGGAGGAGGAGGACAGGACAAAGAAATCACTTGAACCGCAAGTGGAGTTG 26442

QY 680 CAGTGAAGTGAATCGTGCCA-CTCATTCCAACTGGGAGACAGAGTGAACATTTGTCT 738  
|||||

Db 26441 CAGTGAGCCGAGATTGAGGTACCCACTCCAGCCTGGCGACAGAGGAGACTCAATCTC 26382

QY 739 CAAAAAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 798  
|||||

Db 26381 AGAAACAAAAATAAA-----CAAAACAAAAAGCTGATAATGTTAAATTTTAAACAATAAA 26328

QY 799 TGATCCCAATACAAATCAATTTCAAGTGTAAATTAATATAAACAATATGAATGAGATA 858  
|||||

Db 26327 TTTATTTAAACCGTTATATCTAAATGTCAATTTGAACATGTAATCAATGTGAACCTTTTA 26268

QY 859 CTTTACATTTCTTTCTTTGTTTTCATA-TTAAGTCTTTGAAAGTGAGTATATATGTTATGC 917  
|||||

Db 26267 ATGGAATATTTCATTTCTTTTAAAGATAAGTCTTTGAGTCCAGTGTGTTATATACAC 26208

QY 918 TGACAGACACA---TCTCAATTTGGACTAGCTACATTTTCAGTGTCTCAGTAGCCACATGTG 974  
|||||

Db 26207 TTATAGCACAACTTCTCAATTTGGACTAGCCACATTTCAAGTGTCTCAAGGACGACATGTG 26148

QY 975 GCTAGCAGTTACTGTATTGGATGGCA 1000  
|||||

Db 26147 GCCAGTGGCTACTGTTTGGGATAACA 26122

RESULT 8  
ACA60949/c  
ID ACA60949 standard; DNA; 55827 BP.

XX AC

AC ACA60949;

DT 11-AUG-2003 (first entry)

XX DT

DE DNA encoding human carboxypeptidase.

XX DE

KW Human; gene; carboxypeptidase; inflammation; cancer; arteriosclerosis;  
neurodegenerative disease; protease; ds; single nucleotide polymorphism;  
SNP.

XX KW

OS Homo sapiens.

XX OS

PH Key

FT variation

FT Location/Qualifiers

FT replace(859,T)

FT /tag= u

FT /note= "Single nucleotide polymorphism"

FT CDS

FT /product= a

FT /tag= a

FT /product= "Carboxypeptidase"

FT exon

FT 2017..2090

FT /tag= b

FT /number= 1

FT intron

FT 2091..19206

FT /tag= c

FT /number= 1

FT variation

FT /tag= v

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= w

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= x

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= y

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= z

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= aa

FT /note= "Single nucleotide polymorphism"

FT variation

FT /tag= ab

FT /note= "Single nucleotide polymorphism"

FT FT



FT	variation	replace(6254,C)	FT		/*tag= as
FT		/*tag= ac	FT	variation	/note= "Single nucleotide polymorphism"
FT	variation	replace(10171,C)	FT		replace(31600,C)
FT		/*tag= ad	FT		/*tag= at
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(10452,A)	FT		replace(31714,C)
FT		/*tag= ae	FT		/*tag= au
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(11613,A)	FT		replace(31715,T)
FT		/*tag= af	FT		/*tag= av
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(12130,T)	FT		replace(32193,G)
FT		/*tag= ag	FT		/*tag= aw
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(17867. .17869,GA)	FT		replace(32341,A)
FT		/*tag= ah	FT		/*tag= ay
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		replace(18243,G)	FT		replace(32341,G)
FT		/*tag= ai	FT		/*tag= ax
FT	exon	/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT		19207. .19282	FT		replace(32561,G)
FT		/*tag= d	FT		/*tag= az
FT		/number= 2	FT	variation	/note= "Single nucleotide polymorphism"
FT	intron	19283. .22683	FT		replace(32600. .32602,AA)
FT		/*tag= e	FT		/*tag= ba
FT		/number= 2	FT	variation	/note= "Single nucleotide polymorphism"
FT	variation	replace(22450,C)	FT		replace(32642,T)
FT		/*tag= aj	FT		/*tag= bb
FT		/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT	exon	22684. .22808	FT		replace(32793. .32795,TG)
FT		/*tag= f	FT		/*tag= bc
FT		/number= 3	FT	exon	/note= "Single nucleotide polymorphism"
FT	intron	22809. .24477	FT		33053. .33157
FT		/*tag= g	FT		/*tag= l
FT		/number= 3	FT	variation	/number= 6
FT	variation	replace(23003,T)	FT		replace(33071,A)
FT		/*tag= ak	FT		/*tag= bd
FT		/*note= "Single nucleotide polymorphism"	FT	intron	/note= "Single nucleotide polymorphism"
FT	variation	replace(24055. .24057,GA)	FT		33158. .42288
FT		/*tag= al	FT		/*tag= m
FT		replace(24132,T)	FT		/number= 6
FT	variation	/*note= "Single nucleotide polymorphism"	FT	variation	replace(34721,T)
FT		/*tag= am	FT		/*tag= be
FT		24478. .24586	FT	variation	/note= "Single nucleotide polymorphism"
FT	exon	/*note= "Single nucleotide polymorphism"	FT		replace(35304,A)
FT		/*tag= h	FT		/*tag= bf
FT		/number= 4	FT	variation	/note= "Single nucleotide polymorphism"
FT	intron	24587. .28129	FT		replace(35425,C)
FT		/*tag= i	FT		/*tag= bg
FT		/number= 4	FT	variation	/note= "Single nucleotide polymorphism"
FT	variation	replace(25713,A)	FT		replace(36050,G)
FT		/*tag= an	FT		/*tag= bh
FT		/*note= "Single nucleotide polymorphism"	FT	variation	/note= "Single nucleotide polymorphism"
FT	exon	28130. .28231	FT		replace(36291,G)
FT		/*tag= j	FT		/*tag= bi
FT		/number= 5	FT	exon	/note= "Single nucleotide polymorphism"
FT	intron	28232. .33052	FT		42289. .42382
FT		/*tag= k	FT		/*tag= n
FT		/number= 5	FT		/number= 7
FT	variation	replace(28476,C)	FT	intron	42383. .48643
FT		/*tag= ao	FT		/*tag= o
FT		/*note= "Single nucleotide polymorphism"	FT		/number= 7
FT	variation	replace(29404,G)	FT	exon	48644. .48846
FT		/*tag= ap	FT		/*tag= p
FT		/*note= "Single nucleotide polymorphism"	FT		/number= 8
FT	variation	replace(31047,T)	FT	intron	48847. .51176
FT		/*tag= aq	FT		/*tag= q
FT		/*note= "Single nucleotide polymorphism"	FT		/number= 8
FT	variation	replace(31445,C)	FT	exon	51177. .51264
FT		/*tag= ar	FT		/*tag= r
FT		/*note= "Single nucleotide polymorphism"	FT		/number= 9
FT	variation	replace(31447,C)	FT	intron	51265. .53224
FT			FT		/*tag= s



```
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT replace(23003,T)
FT FT /*tag= ak
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(24055..24056,GAA)
FT FT /*tag= al
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(24132,C)
FT FT /*tag= am
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 24478..24586
FT FT /*tag= h
FT FT /number= 4
FT FT 24587..28129
FT FT /*tag= i
FT FT /number= 4
FT FT /cons_splice= (5'site:no,3'site:no)
FT FT replace(25713,A)
FT FT /*tag= an
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 28130..28231
FT FT /*tag= j
FT FT /number= 5
FT FT 28232..33052
FT FT /*tag= k
FT FT /number= 5
FT FT replace(28476,C)
FT FT /*tag= ao
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(29404,G)
FT FT /*tag= ap
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31047,T)
FT FT /*tag= aq
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31445,C)
FT FT /*tag= ar
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31447,C)
FT FT /*tag= as
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31600,C)
FT FT /*tag= at
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31714,C)
FT FT /*tag= au
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(31715,T)
FT FT /*tag= av
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32193,G)
FT FT /*tag= aw
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32341,R)
FT FT /*tag= ax
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32561,G)
FT FT /*tag= ay
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32600..35601,AAA)
FT FT /*tag= az
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32642,T)
FT FT /*tag= ba
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(32793..32794,TAG)
FT FT /*tag= bb
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 33053..33157
FT FT /*tag= 1
FT FT /number= 6
FT FT replace(33071,A)
FT FT /*tag= bc

FT intron /standard_name= "Single nucleotide polymorphism"
FT FT 33158..42288
FT FT /*tag= m
FT FT /number= 6
FT FT /cons_splice= (5'site:no,3'site:no)
FT FT replace(34721,T)
FT FT /*tag= bd
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(35304,A)
FT FT /*tag= be
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(35425,C)
FT FT /*tag= bf
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(36050,G)
FT FT /*tag= bg
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT replace(36291,G)
FT FT /*tag= bh
FT FT /standard_name= "Single nucleotide polymorphism"
FT FT 42289..42382
FT FT /*tag= n
FT FT /number= 7
FT FT 42383..48643
FT FT /*tag= o
FT FT /number= 7
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT 48644..48846
FT FT /*tag= p
FT FT /number= 8
FT FT 48847..51176
FT FT /*tag= q
FT FT /number= 8
FT FT /cons_splice= (5'site:yes,3'site:no)
FT FT 51177..51264
FT FT /*tag= r
FT FT /number= 9
FT FT 51265..53224
FT FT /*tag= s

Query Match 24.0%; Score 240.2; DB 10; Length 55827;
Best Local Similarity 68.6%; Pred. No. 8.6e-42;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

Qy 363 TCATACCTAGAGTAGGTGGTTTGTAGTAGAAATAATGCTGAGCTGCTTATGTCATTTCCTCA 422
Db 46943 TGAATCCTCAGACCAGGGTTGTTTAAAGAAATATAATGCAAGTCACATATTTACTTTAA 46884

Qy 423 GTTTTGTAGTAGGCACATTAAACAGGTAAAA--AAGGCTGGGGCGCAGTGGCTCACACCT 480
Db 46883 ATTTTCAAAGTTACCACCTTTAAGAAAAATAAAAAAGAGCGGCAGGCACGCTGCCTCACACCT 46824

Qy 481 GTATCCCGACACTTTGGGAGGCTTGAGCGAGGATCACCTTTGGTCAGGAGTTTGAGA 540
Db 46823 ATATCCCGACACTTTGGGAGGCTTGAGCGAGGCGGATCCTTGAGATCAGAGTTTGACA 46764

Qy 541 CTAGCCTGGCCACATGCGCAAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCTGGC 600
Db 46763 CCAGCGAGGCGACATGCTGAAACCCGCTCTTACCAAAA--TACAAAAATACCCAGGC 46706

Qy 601 ATGCTGGCGGCGCTCTTAATCTCAGCTGCTCAGAGGCGGAGACACAAGAAATCACTTAA 660
Db 46705 ATAGTGTGCAGCGCTCTTAATCCAGCTACTTTGGGAGGCTGAGGACGCAAACTTTGCTTGA 46646

Qy 661 ACCGAGAGGTGGAGGTTTGCAGTGAGCTGCTGCCACT-CACCTCCACCTGGGAGA 719
Db 46645 ACCGAGAGGCGAGGTTTGCAGTGAGCTGAGATCATGACACTGCACCTCCAGCTGGAGTGA 46586

Qy 720 CAGAGTGACA-CTTTTGTCTCAAAAGAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
Db 46585 CAGAACAGACCCCTGCCACACACACACACACAAAAAGAAATATGTGAGATTAATTCAATA 46526

Qy 779 AGTTAACTTTTAATAACCAATGTATCCCAATAACAATCAATTTCAAGTGTAAATAATA 838
```

46525 TTGTAATTTTAAACCAATATGTACAAAGTATCA-----TTTCAGTGTACCAATATA 46470  
839 AAACAATATGATGAGATCTTTACATCTCTTTCTTTTTCATPATTAAGTCTTTGAAA 898  
46469 AA--AAATATTCTTGAATAATTTTATGTACTT-----TTTTCATPACTAAGCCTTCAAT 46417  
899 GTGAGTATATGTTATGCTGACACACATCTCAATTTGGACTAGCTACATTTTCAGGTGC 958  
46416 TCTGTTGTATTTAACTACTGATGTATCTCAATTCAGACAAGCTACATTTTAAGCAC 46357  
959 TCAGTAGCCACATGTGGCTAGCAGTTACTGTATTGGA 995  
46356 TGGATAGCCACATGTGGTTAGTGGCTACTATGCTGGA 46320

## RESULT 10

ID ADS36454  
AD ADS36454 standard; DNA; 58337 BP.

AC ADS36454;

XX  
DT 16-DEC-2004 (first entry)

DE Human autoimmune disease-related genomic DNA sequence - SEQ ID 1668.

XX single nucleotide polymorphism detection; SNP detection;  
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;  
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;  
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;  
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;  
KW primary systemic vasculitis; ds.

XX Homo sapiens.

OS  
XX WO2004083403-A2.

XX 30-SEP-2004.

XX 18-MAR-2004; 2004WO-US008461.

XX 18-MAR-2003; 2003US-0455444P.

PR 25-APR-2003; 2003US-0465241P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Begovich AB, Alexander HC;

XX WPI; 2004-728480/71.

XX New isolated nucleic acid molecule comprises at least 8 contiguous  
PT nucleotides where one of the nucleotides is a single nucleotide  
PT polymorphism (SNP), useful for diagnosing or treating autoimmune  
PT diseases, e.g. rheumatoid arthritis.

PS Claim 16; SEQ ID NO 1668; 123pp; English.

XX The invention comprises amino acid and coding sequences containing  
CC genetic polymorphisms associated with an altered risk of developing an  
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further  
CC comprises a method of identifying an individual that has an altered risk  
CC of developing an autoimmune disease, comprising detecting a single  
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA  
CC and protein sequences of the invention are useful for diagnosing and  
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1  
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory  
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious  
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,  
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The  
CC present nucleic acid represents a human autoimmune disease-related  
CC genomic DNA sequence of the invention. NOTE: The present sequence is not  
CC shown in the specification, but has been retrieved from the WIPO website.

XX

SQ Sequence 58337 BP; 18163 A; 11254 C; 11433 G; 17380 T; 0 U; 107 Other;  
Query Match 24.0%; Score 240.2; DB 13; Length 58337;  
Best Local Similarity 68.6%; Pred. No. 8.7e-42;  
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;  
QY 363 TCATACCTAGAGTAGTGGTGTGTTAGTAGAAATAATGCTGAGCTGCTTATGTCATTTCCA 422  
DB 13209 TGAATCTTCAGACGAGGGTGTGTTTAAAGAAATATATGCAAGTCACATATTTACTTTAA 13268  
QY 423 GTTTTTTAGTAGGCACATTAATAACAGGTAAAA--AAGGCTGGGGCGCAGTGGCTCACACCT 480  
DB 13269 ATTTTCAAGTTACCACCTTTAAGAAAAATAAAAGAGCCAGGCACGCTGGCTCACACCT 13328  
QY 481 GTAATCCAGCAGCTTTGGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGAGTTTGAGA 540  
DB 13329 ATAATCCAGCAGCTTTGGGAGGCTGAGCGAGGAGATCACTTGAGATCAGGAGTTTGACA 13388  
QY 541 CTAGCCTGGCCCAACATGCGGAAACTCTGTCTCTAAAAAAAATAACAAAAATTAGCCTGGC 600  
DB 13389 CCAGCGAGGCCAATGTTGAAACCCCGTCTCTACCAAAA--TACAAAAATAGCCAGGC 13446  
QY 601 ATGCTGGCGGCGCCTGTAAATCTCAGCTGCTCAGGAGGCCGAGACACAAAGAAATCACTTAA 660  
DB 13447 ATAGTGTGCGAGCGCTCTGAATCCAGCTACTTTGGGAGGCTGAGGCACGAAACTTGTGTA 13506  
QY 661 ACCAGAGGTGGAGGTTGAGTGAGTGAGTGCCTCACT--CACTCCAACTGGGAGAGA 719  
DB 13507 ACCAGGAGGCGAGGTTGCGAGTGAGTGAGTGCATGACACTGCCTCCAGCTGGATGA 13566  
QY 720 CAGAGTGACA-CTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAGGTGA 778  
DB 13567 CAGAACAGACCTTGCACACACACACACAAAGAAATATGTGAGATTAAATTCATA 13626  
QY 779 AGTTAACTTTAATAACCAATGTATCCCAATACAATCATTTCAAAGTGAATTAATATA 838  
DB 13627 TTGTAATTTATTAAACCAATATGTACAAGTATCA---TTTCAAGTGAACCAATATA 13682  
QY 839 AAACAATATGATGAGATCTTTTACATCTTTTCTTTGTTTTCATATTAAGTCTTTGAAA 898  
DB 13683 AA--AATTAATCTTGAATAATTTTATGTACTT-----TTTTCATACTAAGCCTTCAAT 13735  
QY 899 GTGAGTATATGTTATGCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGGTGC 958  
DB 13736 TCTGGTGTGTTATTAAACACTCAGTGTATGTCTCAATTCAGACAAGCTACATTTTAAGCAC 13795  
QY 959 TCAGTAGCCACATGTGCTAGCAGTTACTGTATTGGA 995  
DB 13796 TGGATAGCCACATGTGGTTAGTGGCTACTATGCTGGA 13832

## RESULT 11

ADS36462/c

ID ADS36462 standard; DNA; 64423 BP.

XX ADS36462;

XX 16-DEC-2004 (first entry)

DT Human autoimmune disease-related genomic DNA sequence - SEQ ID 1676.

DE single nucleotide polymorphism detection; SNP detection;  
XX rheumatoid arthritis; type 1 diabetes; multiple sclerosis;  
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;  
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;  
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;  
KW primary systemic vasculitis; ds.

XX Homo sapiens.

OS  
XX WO2004083403-A2.

XX 30-SEP-2004.



```
Db 11770 GAGCCGAGGTGGGCGAGATCACCTGAGGTGAGGAGTTTGAGACCAGCCTGGCCCAACATGG 11711
QY 559 CGAACTCTGCTCTAATAAATAAATAGCTGGATGGTGGGGGCGCCTGT 618
Db 11710 TGAACCCCGTCTCT-ACTAAGATATATAAATAATAGCTGGGAGTGGTGGTGGCCTGT 11652
QY 619 AATCTCAGCTGCTCAGGAGGCGGAGACACAGAATCACTTAAACCCAGGAGGTGGAGTT 678
Db 11651 AATCCAGCTACTCAGAGGCTGAGGAGGAGAAATGCTTGAACCCAGGAGGCGAGAGTT 11592
QY 679 GCACTGAGCTGAGTCTGCGCACTCACTCCCACTGGGAGACAGAGTGACACTTTTGTCT 738
Db 11591 GCACTGAGCAGAGCTGCACCATTCCTCCAGCATGAGTGCAGGCGGAAAGTCTGT 11532
QY 739 CAAAAGAAAAAATAAACAAGTAAAAAAGAAACAGGTGAAGTTAACTTTAATAACCAA 798
Db 11531 CCAAAAAAATAAATAAATAA-----AAAAAGCAACAGCTTTTGGGCAATGATCACTACATTAC 11477
QY 799 TGTATCCCAATACATCAATTTCAAAGTGTAATTAATAAATAAACAATATGAATGAGATA 858
Db 11476 TACACCTCCCTCCCTTCATTTTCAAAAAAATTTTCAATGCAATTAACAACATTAATA 11417
QY 859 CTTTACATTTCTTTCTGTTTTTCAATTAAGTCTTTGAAAGTGAGTATATATGTTATGC- 917
Db 11416 GATTACATTTCT---TTCTTCTCATGTTAAGTCTTTGAAATCTGGGGTATATTTATACT 11361
QY 918 TGACAGCACATCTCAATTTGGACTAGCTACATTTTCAGTGTCTCAGTAGCCACATGTGGCT 977
Db 11360 TAAAGCATGCTCTAAATTTGGACTAGACACATTTCAAGTGTCTCAGTAGCTACAGTGGAT 11301
QY 978 AGCAGTACTCTATTGGATG 997
Db 11300 AGTGCTACTGTATTAGATG 11281

RESULT 13
AAK68359/c
ID AAK68359 standard; DNA; 2953 BP.
XX
AC AAK68359;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23171.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD
XX
PP 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215133P.
PR 07-JUL-2000; 2000US-0215647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
```





PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0198874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225259P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234597P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 20-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
PR PA (HUMA-) HUMAN GENOME SCI INC.  
PR PI Rosen CA, Barash SC, Ruben SM;  
PR XX WPI; 2001-483426/52.  
PR XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PR PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
PR XX Disclosure; SEQ ID NO 23172; 3071pp + Sequence Listing; English.  
PR XX

```
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK64702 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;

Query Match      22.7%; Score 227.4; DB 4; Length 2953;
Best Local Similarity 69.1%; Pred. No. 2.9e-39;
Matches 403; Conservative 0; Mismatches 161; Indels 19; Gaps 6;

Qy 414 TCATTTCCAGTTTTTTAGTAGCCACATTAAACAGGTAATAAAA-----GGCTGGGCGCA 467
Db 1699 TAATTAAGATTTTTTAACAGCCACATTATTAATGTAAGAAAAAAGTCGGCGCGCG 1640

Qy 468 GTGGCTCACCTGTATCCAGACACTTTGGGAGGCTGAGGAGCGAGATCACCTTTGGT 527
Db 1639 GTGGCTCACGCTGTATCCAGACACTTTGGGAGACCGAGGAGCGGATCAC--GAGCT 1582

Qy 528 CAGGAGTTTGAGACTAGCTCGCCAAACATGCGAACTCTCTCTAATAAATAACAA 587
Db 1581 CGAAGATCGAGACCATCTTTGTAACACGCGTGAACCTCCGCTCTACTAATAAATAACAA 1522

Qy 588 AAATTAGCTGGCATGTGGGGCGGCTGTAATCTCAGCTGCTCAGAGGCGCGAGACAC 647
Db 1521 AAATTAGCGGCGATGTGGCGGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGAG 1462

Qy 648 AAGAAATCACTTAAACCCAGAGGTGGAGTTGCGAGTGAGCTGAGATCGTCCACTCACTC 707
Db 1461 GAGAATGGCGTCAACCCGAGGAGAGAGCTTGCAGTGAGCTGAGATGAGATGAGTCACTGCACT 1402

Qy 708 CAACCTGGGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAGGTAATAAAA 767
Db 1401 CCACCAGCTGGGCGACAGAGAC--AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAAA 1343

Qy 768 GAAACAGGTGAAGTTAACTTTAATAACCAATGATATCCAAATACAATCAATTTCAAGTG 827
Db 1342 TAAGTCATAGATTTTAATATATTAGTCTCAGTATGCTTAAATATATATCATTTAA----- 1288

Qy 828 TAATTAATATAAACAATATGAATGAGATCTTTACATTTCTTTTCTTTTTCATATTA 887
Db 1287 -CATGTAATCAACAAGTTATTAACAATAATTTCTACATTTCTTTGTTGGGGTAA--A 1232

Qy 888 AGTCTTTGAAGTGAATATATG--TTATGCTGACAGACATCTCAATTTGGACTAGCTA 946
Db 1231 AATCTTTGAAATTAATATGTTATTTTATCTTACAGCTGATCTGTGATTCAGACTAGCCA 1172

Qy 947 CATTTTCAGTGTCTAGTAGCCACATGTGGCTAGCAGTTACTGT 989
Db 1171 CATTTTCAGTGTCTAGTAGCCCATGTGGCTAGCACTATGTT 1129

RESULT 15
ADQ64563/c
ID ADQ64563 standard; cDNA; 3377 BP.
XX
AC ADQ64563;
XX
DT 07-OCT-2004 (first entry)
XX
```

```
DE Novel human cDNA sequence #1724.
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
XX cytosstatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
XX neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
XX cancer.
XX Homo sapiens.
XX EPI440981-A2.
XX 28-JUL-2004.
XX 21-JAN-2004; 2004EP-00001196.
XX 21-JAN-2003; 2003JP-00102206.
XX 09-MAY-2003; 2003JP-00131392.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
XX Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
XX P-PSDB; ADQ66751.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
XX Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 1724; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
XX polypeptides, sequences hybridizing to these nucleotides, sequences
XX encoding partial polypeptides and sequences having 70% or 90% identity to
XX the nucleotide and protein sequences. The nucleotides and polypeptides
XX are useful as diagnostic markers or therapeutic target for the diseases
XX or morbid states. They are also useful for treating osteoporosis,
XX neurological diseases, Alzheimer's diseases, Parkinson's diseases,
XX dementia and various cancers. This sequence corresponds to a nucleotide
XX sequence of the invention.
XX Sequence 3377 BP; 937 A; 660 C; 718 G; 1062 T; 0 U; 0 Other;

Query Match      22.7%; Score 227.4; DB 12; Length 3377;
Best Local Similarity 69.1%; Pred. No. 2.9e-39;
Matches 403; Conservative 0; Mismatches 161; Indels 19; Gaps 6;

Qy 414 TCATTTCCAGTTTTTTAGTAGCCACATTAAACAGGTAATAAAA-----GGCTGGGCGCA 467
Db 2132 TAATTAAGATTTTTTAACAGCCACATTATAATGTAAGAAAAAAGTCGGCGCGCG 2073

Qy 468 GTGGCTCACACTGTATCCAGACACTTTGGGAGGCTGAGGAGCGAGATCACCTTTGGT 527
Db 2072 GTGGCTCACGCTGTATCCAGACACTTTGGGAGACCGGAGCGGAGTATCAC--GAGCT 2015

Qy 528 CAGGAGTTTGAGACTAGCTGCGCAACATGCGAACTCTCTCTAATAAATAACAA 587
Db 2014 CGAAGATCGAGACCATCTTTGTAACCGGTGAATCTCCGCTCTACTAATAAATAACAA 1955

Qy 588 AAATTAGCTGGCATGTGGGGCGGCTGTAATCTCAGCTGCTCAGAGGCGCGAGACAC 647
Db 1954 AAATTAGCGGCGATGTGGCGGATGCTGTAGTCCAGCTACTTTGGAGGCTGAGGCGAG 1895

Qy 648 AAGAAATCACTTAAACCCAGAGGTGGAGTTGCGAGTGAGCTGAGATCGTCCACTCACTC 707
Db 1894 GAGAATGGCGTGAACCCGAGGAGGAGAGCTTGCAGTGAGCTGAGATGAGTATGAGTATG 1835

Qy 708 CAACCTGGGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAGGTAATAAAA 767
Db 1834 CCACGAGCTGGGCGAGAGAGAC--AGACTCTGTCTCAGAAAAAAGAAAAAAGAAAAA 1776

Qy 768 GAAACAGGTGAAGTTAACTTTAATAACCAATGATATCCAAATACAATCAATTTCAAGTG 827
```

```
Db 1775 TAAGTCATAGATTTTAAATATATTTTAGCTCAGTATGCTAAATAATATATCAATTTAA----- 1721
Qy 828 TAAATTAATATAAAACAATATGCAATGAGATACCTTTACATTCCTTTCTGTTTTCATATTA 887
Db 1720 -CATGTAATCAACAAGTTATTACAAATATTTACATTCCTTGTGGGGTAAT---A 1665
Qy 888 AGTCTTTGAAAGTGAGTATATATG-TTATGCTGACAGCACATCTCAATTTGGACTAGCTA 946
Db 1664 AATCTTTGAAATTAATGTAATTTTATACCTTACAGCTGATCTTGATTCAGACTAGCCA 1605
Qy 947 CATTTCCAGTGCTCAGTAGCCACATGTGGCTAGCAGTTACTGT 989
Db 1604 CATTTCCAGTGCTCAGTAGCCCATGTGGCTAGCAACTATGTT 1562
```

Search completed: July 4, 2005, 03:03:25  
Job time : 358.271 secs

**THIS PAGE BLANK (USPTO)**



Db 3545 GTTGCCAAATCTGAGAAATCCAGAAATGCGCAAGACAGAGTCCAGACCCCTGTCTTCACAGA 3604  
Qy 361 GCTCATACCTAGAGTAGTGGTGTAGTAAATATGCTGAGCTGCTTATGTCAATTC 420  
Db 3605 GCTCATACCTAGAGTAGTGGTGTAGTAAATATGCTGAGCTGCTTATGTCAATTC 3664  
Qy 421 CAGTTTTTTAGTACCCACATTAATAACAGGTAAATAAGGCTGGCGCAGTGGCTCACACCT 480  
Db 3665 CAGTTTTTTAGTACCCACATTAATAACAGGTAAATAAGGCTGGCGCAGTGGCTCACACCT 3724  
Qy 481 GTAATCCAGACATTTTGGGAGGCTGAGCGAGCAGATCACTTTTGGTCAGAGTTTGAGA 540  
Db 3725 GTAATCCAGACATTTTGGGAGGCTGAGCGAGCAGATCACTTTTGGTCAGAGTTTGAGA 3784  
Qy 541 CTAGCCTGGCCACATCGCGAACTCTGTCTTAAATAAATAACAAATTAAGCCTGGC 600  
Db 3785 CTAGCCTGGCCACATCGCGAACTCTGTCTTAAATAAATAACAAATTAAGCCTGGC 3844  
Qy 601 ATGTGGCGGCGCTGTAAATCTCAGCTGTCTCAGAGCGCGAGACACAAGAATCACCTAA 660  
Db 3845 ATGTGGCGGCGCTGTAAATCTCAGCTGTCTCAGAGCGCGAGACACAAGAATCACCTAA 3904  
Qy 661 ACCAGAGGTGGAGGTTCAGTGTGAGTGTGCTGCTCACTCCAACTGGGAGAC 720  
Db 3905 ACCAGAGGTGGAGGTTCAGTGTGAGTGTGCTGCTCACTCCAACTGGGAGAC 3964  
Qy 721 AGAGTGACATTTTGTCTCAAAAGAAATAAATAAACAAGTAAATAAACAAGTCAAG 780  
Db 3965 AGAGTGACATTTTGTCTCAAAAGAAATAAATAAACAAGTAAATAAACAAGTCAAG 4024  
Qy 781 TTAACCTTAAATACCCCAATGTATCCCAATAACAATCAATTTCAAAAGTGTAAATATAATAA 840  
Db 4025 TTAACCTTAAATACCCCAATGTATCCCAATAACAATCAATTTCAAAAGTGTAAATATAA 4084  
Qy 841 ACAATTAATGATGAGATACATTTTCAATTTCTTTTCAATTAAGTCTTTGAAAGT 900  
Db 4085 ACAATTAATGATGAGATACATTTTCAATTTCTTTTCAATTAAGTCTTTGAAAGT 4144  
Qy 901 GAGTATATATTTATGCTGACAGACATCTCAATTTGAGTGTAGTACATTTTCAGGTCTC 960  
Db 4145 GAGTATATATTTATGCTGACAGACATCTCAATTTGAGTGTAGTACATTTTCAGGTCTC 4204  
Qy 961 AGTAGCCACATGTGGCTAGCAGTGTACTGTATTGATGGCAC 1001  
Db 4205 AGTAGCCACATGTGGCTAGCAGTGTACTGTATTGATGGCAC 4245

RESULT 2  
US-09-949-016-63399/c  
; Sequence 63399, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63399  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63399

Query Match 60.0%; Score 600.6; DB 4; Length 601;  
Best Local Similarity 99.8%; Pred. No. 3.2e-148;  
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 148 TAAGTCAGAAAAGGTCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 207  
Db 601 TAAGTCAGAAAAGGTCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCG 542  
Qy 208 CCAGTCACAAAAGCAGGTAGAGACTCTCTCCAGGTGAGGACACCTCTCTTTATTTCAGCAGA 267  
Db 541 CCAGTCACAAAAGCAGGTAGAGACTCTCTCCAGGTGAGGACACCTCTCTTTATTTCAGCAGA 482  
Qy 268 TACACACTGAGTCCCAACTCGGTAAATGAGAGCGTTGCCAAATTCGAGAAATCAGCAAT 327  
Db 481 TACACACTGAGTCCCAACTCGGTAAATGAGAGCGTTGCCAAATTCGAGAAATCAGCAAT 422  
Qy 328 TGCCAAAGACAGTCAGGACCCCTCTCTCACAGAGCTCATACCCCTAGAGTAGTGTGTTTA 387  
Db 421 TGCCAAAGACAGTCAGGACCCCTCTCTCACAGAGCTCATACCCCTAGAGTAGTGTGTTTA 362  
Qy 388 GTAGAAATTAATGCTGAGCTGCTTATGTCAATTTCCAGTTTTTTTAGTAGCCACATTAATAACA 447  
Db 361 GTAGAAATTAATGCTGAGCTGCTTATGTCAATTTCCAGTTTTTTTAGTAGCCACATTAATAACA 302  
Qy 448 GGTAAAAAAGGCTGGCGCAGTGGCTCACACCTGTAAATCCAGACATTTTGGAGGCTGAG 507  
Db 301 RGTAATAAAGGCTGGCGCAGTGGCTCACACCTGTAAATCCAGACATTTTGGAGGCTGAG 242  
Qy 508 GCAGGCAGATCACTTTTGGTCAGAGTTTGAGACTAGCTGGCCCAACATGCGGAACTCT 567  
Db 241 GCAGGCAGATCACTTTTGGTCAGAGTTTGAGACTAGCTGGCCCAACATGCGGAACTCT 182  
Qy 568 GTCTCTAAAAAATAACAAAAATTAGCCTGGCATGGTGGCGGCGCTGTAAATCTCAGC 627  
Db 181 GTCTCTAAAAAATAACAAAAATTAGCCTGGCATGGTGGCGGCGCTGTAAATCTCAGC 122  
Qy 628 TGCTCAGGAGCGCCAGACACAAGAATCACTTAAACCAGAGGTGGAGTTGCAAGTGCAGC 687  
Db 121 TGCTCAGGAGCGCCAGACACAAGAATCACTTAAACCAGAGGTGGAGTTGCAAGTGCAGC 62  
Qy 688 TGAGATCGTCCACTCACTCCAACTGGGAGACAGAGTGACATTTTGTCTCAAAAGAA 747  
Db 61 TGAGATCGTCCACTCACTCCAACTGGGAGACAGAGTGACATTTTGTCTCAAAAGAA 2  
Qy 748 A 748  
Db 1 A 1

RESULT 3  
US-09-949-016-63400/c  
; Sequence 63400, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63400  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63400

Query Match 60.0%; Score 600.6; DB 4; Length 601;  
Best Local Similarity 99.8%; Pred. No. 3.2e-148;  
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 167 GGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAG 226  
Db 601 GGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAG 542

Qy 227 AGACTCTCCGAGTGGAGACACCTCTTTTATTCAGCAGATACACACTGAGTGCACACT 286  
Db 541 AGACTCTCCGAGTGGAGACACCTCTTTTATTCAGCAGATACACACTGAGTGCACACT 482

Qy 287 CGGTAAACATGAGCGTTCGCAAAATCTCAGAAATCCAGCAATTCGCAAGACAGTCCAGACC 346  
Db 481 CGGTAAACATGAGCGTTCGCAAAATCTCAGAAATTCGCAAGACAGTCCAGACC 422

Qy 347 CCTGTCTCAGACAGCTCATACCCCTAGATAGTGGTGTATAGTAAATATGCTGAGCT 406  
Db 421 CCTGTCTCAGACAGCTCATACCCCTAGATAGTGGTGTATAGTAAATATGCTGAGCT 362

Qy 407 GCTTATGTCATTTCCAGTTTTTTTAGTACCCACATTAACAGGTAAAGAGCTGGGCGC 466  
Db 361 GCTTATGTCATTTCCAGTTTTTTTAGTACCCACATTAACAGGTAAAGAGCTGGGCGC 302

Qy 467 AGTGGCTCACACCTGTATCCAGCACTTTGGGAGGCTGAGCGGCGAGATCAGCTTTGG 526  
Db 301 RGTGGCTCACACCTGTATCCAGCACTTTGGGAGGCTGAGCGGCGAGATCAGCTTTGG 242

Qy 527 TCAGGAGTTTGAGACTAGCTGGCGCAACATGGCGAACTCTGTCTCTAAAAAAATACA 586  
Db 241 TCAGGAGTTTGAGACTAGCTGGCGCAACATGGCGAACTCTGTCTCTAAAAAAATACA 182

Qy 587 AAAATTAGCTGGCATGTGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACA 646  
Db 181 AAAATTAGCTGGCATGTGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACA 122

Qy 647 CAAGAATCACTTAACCCAGAGGTGGAGGTTGAGTGAGTGCAGTGCAGTGCCTCACT 706  
Db 121 CAAGAATCACTTAACCCAGAGGTGGAGGTTGAGTGAGTGCAGTGCAGTGCCTCACT 62

Qy 707 CCAACCTGGGAGACAGTGCACCTTTTGTCTCAAAAAGAAAAAACAAGTAAAAA 766  
Db 61 CCAACCTGGGAGACAGTGCACCTTTTGTCTCAAAAAGAAAAAACAAGTAAAAA 2

Qy 767 A 767  
Db 1 A 1

RESULT 4  
US-09-949-016-63401/c  
; Sequence 63401, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63401  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human

US-09-949-016-63401

Query Match 60.0%; Score 600.6; DB 4; Length 601;  
Best Local Similarity 99.8%; Pred. No. 3.2e-148;  
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 242 GAGGACACCTCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAACTGGAGCG 301  
Db 601 GAGGACACCTCTCTTTTATTCAGCAGATACACACTGAGTGCCAACTCGGTAACTGGAGCG 542

Qy 302 TTGCCAAATTCAGAAATCCAGCAATTCAGCAAGACAGTCCAGGACCCCTGTTCTCACAGAG 361  
Db 541 TTGCCAAATTCAGAAATCCAGCAATTCAGCAAGACAGTCCAGGACCCCTGTTCTCACAGAG 482

Qy 362 CTCTATACCTAGAGTAGTGGTGTATAGTAAATTAATGCTGAGTGTCTTATGCTATTC 421  
Db 481 CTCTATACCTAGAGTAGTGGTGTATAGTAAATTAATGCTGAGTGTCTTATGCTATTC 422

Qy 422 AGTTTTTTTAGTAGCCATTAACACAGGTAAAAAGGCTGGCGCAGTGGCTCACACCTG 481  
Db 421 AGTTTTTTTAGTAGCCATTAACACAGGTAAAAAGGCTGGCGCAGTGGCTCACACCTG 362

Qy 482 TAATCCAGCACTTTGGGAGGCTGAGGCGAGATCACCTTTTGGTCAGGAGTTTGAGAC 541  
Db 361 TAATCCAGCACTTTGGGAGGCTGAGGCGAGATCACCTTTTGGTCAGGAGTTTGAGAC 302

Qy 542 TAGCTGGCCAAATGGGAAACTCTGTCTCTAAAAAAATACAATAATAGCTGGCA 601  
Db 301 YAGCTGGCCAAATGGGAAACTCTGTCTCTAAAAAAATACAATAATAGCTGGCA 242

Qy 602 TGGTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACACAGAACTCACTTAAA 661  
Db 241 TGGTGGCGGCGCTGTAATCTCAGCTGCTCAGGAGCGCGAGACACAGAACTCACTTAAA 182

Qy 662 CCCAGGAGGTGGAGTTGCGAGTGCAGATCGTCCACTCTCAACCTGGGAGACA 721  
Db 181 CCCAGGAGGTGGAGTTGCGAGTGCAGTGCAGATCGTCCACTCTCAACCTGGGAGACA 122

Qy 722 GAGTGACACTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAGAACAGGTGAAGT 781  
Db 121 GAGTGACACTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAGAACAGGTGAAGT 62

Qy 782 TAACCTTTAATACCAATGTATCCCAATACAACTCAATCTCAAGTGAATTAATATAA 841  
Db 61 TAACCTTTAATACCAATGTATCCCAATACAACTCAATCTCAAGTGAATTAATATAA 2

Qy 842 C 842  
Db 1 C 1

RESULT 5  
US-09-949-016-63402/c  
; Sequence 63402, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63402  
; LENGTH: 601  
; TYPE: DNA



; ORGANISM: Human  
US-09-949-016-63402

```
Query Match      54.5%; Score 546; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.1e-134;
Matches 546; Conservative 0; Mismatches 0; Indels
```

Qy	456	AGGCTGGCGCAGTGGCTCACACCTGTAAATCCAGCACTTTGGGGAGGCTGAGGCAGGCAG	515
Db	601	AGGCTGGCGCAGTGGCTCACACCTGTAAATCCAGCACTTTGGGGAGGCTGAGGCAGGCAG	542
Qy	516	ATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCCAACATGGCGAAACTCTGTCTCTAA	575
Db	541	ATCACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCCAACATGGCGAAACTCTGTCTCTAA	482
Qy	576	AAAAAATACAAAAATTAGCCTGGCATGCTGGCGGGCGCCTGTAAATCTCAGCTGCTCAGG	635
Db	481	AAAAAATACAAAAATTAGCCTGGCATGCTGGCGGGCGCCTGTAAATCTCAGCTGCTCAGG	422
Qy	636	AGGCCGAGACACAAGATTCATTTAAACCCAGAGAGTGAGGTTGCAGTGAGCTGAGATCG	695
Db	421	AGGCCGAGACACAAGATTCATTTAAACCCAGAGAGTGAGGTTGCAGTGAGCTGAGATCG	362
Qy	696	TGCCACTCCTCAACCTGGCGAGACAGAGTGACACTTTTGTCTCAAAAGAAAAA	755
Db	361	TGCCACTCCTCAACCTGGCGAGACAGAGTGACACTTTTGTCTCAAAAGAAAAA	302
Qy	756	ACAAGTAAAAAGAAACAGGTGAAGTTAACTTTTAATAACCCAAATGATCCCAATACAAAT	815
Db	301	ACAAGTAAAAAGAAACAGGTGAAGTTAACTTTTAATAACCCAAATGATCCCAATACAAAT	242
Qy	816	CATTTCAAAGTGTAATTAATAATAAAACAATTTATGAATGAGATACITTTACATCTTTTCTT	875
Db	241	CATTTCAAAGTGTAATTAATAATAAAACAATTTATGAATGAGATACITTTACATCTTTTCTT	182
Qy	876	GTTTTCATATTAAGCTTTGAAAGTGAGTATATATGTTATGCTGCACAGCATCTCAATTT	935
Db	181	GTTTTCATATTAAGCTTTGAAAGTGAGTATATATGTTATGCTGCACAGCATCTCAATTT	122
Qy	936	TGGACTAGCTACATTTCAAGTGCCTCAGTAGGCACATGTGGCTAGCAGTTACTGTATTGGA	995
Db	121	TGGACTAGCTACATTTCAAGTGCCTCAGTAGGCACATGTGGCTAGCAGTTACTGTATTGGA	62
Qy	996	TGGCAC 1001	
Db	61	TGGCAC 56	

RESULT 6  
US-09-949-016-63403/c  
; Sequence 63403, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,769  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63403  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63403

Query Match	25.4%;	Score 254;	DB 4;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 3.9e-57;		

Qy	748	AAAAAAAAACAGTAAAAAGAAACAGGTGAAGTTAACTTTAAATAACCCAAATGTATCCCA	807
Db	601	AAAAAAAAACAGTAAAAAGAAACAGGTGAAGTTAACTTTAAATAACCCAAATGTATCCCA	542
Qy	808	AATACAATCATTTCAAAGTGTAAATTAATAATAAAACAATTATGAATGAGATACTTTACATT	867
Db	541	AATACAATCATTTCAAAGTGTAAATTAATAATAAAACAATTATGAATGAGATACTTTACATT	482
Qy	868	CTTTCTCTGTTTTTCATATTAAGTCTTTGAAAAGTCAGTATATATGTTATGCTGACAGCACA	927
Db	481	CTTTCTCTGTTTTTCATATTAAGTCTTTGAAAAGTCAGTATATATGTTATGCTGACAGCACA	422
Qy	928	TCTCAATTTGGACTAGCTACATTTCAAGTGTCTCAGTAGCCACATGTGGCTAGCAGTTACT	987
Db	421	TCTCAATTTGGACTAGCTACATTTCAAGTGTCTCAGTAGCCACATGTGGCTAGCAGTTACT	362
Qy	988	GTATTGGATGGCAC	1001
Db	361	GTATTGGATGGCAC	348

RESULT 7  
 US-09-813-133A-3/c  
 ; Sequence 3, Application US/09813133A  
 ; Patent No. 6455294  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GAN, Weiniu et al  
 ; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
 ; TITLE OF INVENTION: USES THEREOF  
 ; FILE REFERENCE: CL001173  
 ; CURRENT APPLICATION NUMBER: US/09/813,133A  
 ; CURRENT FILING DATE: 2001-06-06  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 55827  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-813-133A-3

Query Match	24.0%;	Score 240.2;	DB 3;	Length 55827;
Best Local Similarity	68.6%;	Pred. No. 1.2e-52;		
Matches 437;	Conservative 0;	Mismatches 183;	Indels 17;	Gaps 7;
Qy	363	TCATACCCCTAGAGTAGTGCTGTTAGTAGAANAATATGTGAGCTGCTTATGTCTCATTTCCA	422	
Db	46943	TGAATCCTCAGACCAGGGTGTTTAATAGAAATATAATGCAAGTCACATATTTTACTTTTAA	46884	
Qy	423	GTTTTTTTAGTAGCCACATTAAAAACAGGTAAAA--AAGGCTGGCGGCAGTGGCTCACACCT	480	
Db	48883	ATTTTCAAGTACCATTTTAGAAAAATAAAAAGAGGCCAGCGTGGCTCACACCT	46824	
Qy	481	GTAATCCCAGCACTTTTGGAGGCTGAGCGAGCGAGATCACCTTTGGTCAGAGATTTTGAGA	540	
Db	46823	ATAATCCCAGCACTTTTGGAGGCTGAGCGAGCGGATCACTTGAGATCAGGAGATTTTGACA	46764	
Qy	541	CTAGCCTGGCCACATGGCGAAACTCTGTCTCTTAAAAAAAATACAAAAATTTAGCCTGGC	600	
Db	46763	CCAGCAGGCCCAACATGTTGTAACCCCGTCTCTACCAAAA--TACAAAAATTAGCCAGGC	46706	
Qy	601	ATGGTGGCGGGCGCCTGTAAATCTCAGCTGCTCAGGAGGCCGAGACACAAGAAATCACTTAA	660	
Db	46705	ATAAGTGGTGCA CGCCTGTGTAATCCAGCTACTTTGGGAGGCTGAGGCACGNAACTTCTTGA	46646	
Qy	661	ACCCAGGAGGTGGAGTTTGCAGTAGCTGAGATCGTGGCACT-CACTCCAACTGGGAGA	719	
Db	46645	ACCCAGGAGCGAGAGTTTGCAGTAGCTGAGATCATGACATCTGCACTCCAGCCTGGATGA	46586	



Db 64131 AACCCGGAGCGAGAGTTGTCAGTGAACCGAGATCGCGCCACTGCACTCTAGCTGGGTG 64072  
Qy 719 ACAGAGTGACACTTT-TGTCTCAAAAGAGAAAAAACAAGTAAAGAAACACAGGTG 777  
Db 64071 ACAGAGAGAGACTACATCTCAACAAACAACAACAACAATAAATAATATAAGAGA 64012  
Qy 778 AAGTTAACTTTTAAFAACCCCAATGTATCCAAATACAATCAATTTTCAAAAGTGTAAATTAAT 837  
Db 64011 TAGGTAGAAATGATTTTATGTTTAT--CTAACCAATATATCAAAATATTTCAATTCA 63954  
Qy 838 AAAACAATATGAATGAGATACATTCTTTTCTTTTCTTTTCAATTAAGTCTTTGAA 897  
Db 63953 ACATGTAATCAATATAAAAAAATAAGATACCTTTATATCTTTTTCACACTAAGATTTGA 63894  
Qy 898 AGTGAGTATATGTTATGTCGACAGCACATCTCAATTTTGACTAGCTAGCTACATTTCCAGGTG 957  
Db 63893 AATGGTATATATTTTACACTTATAGCATATCTCAATTCAGAAATAGCCTCATAGTGGTG 63834  
Qy 958 CTCA 961  
Db 63833 GGCA 63830

RESULT 10  
US-09-949-016-15779/c  
; Sequence 15779, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15779  
; LENGTH: 265038  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(265038)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15779

Query Match 23.0%; Score 230.2; DB 4; Length 265038;  
Best Local Similarity 69.1%; Pred. No. 1e-49;  
Matches 375; Conservative 0; Mismatches 158; Indels 10; Gaps 4;  
Qy 437 ACATTAACAGGTAAAGGCTGGGCGAGTGGCTCACACCTGTAAATCCAGCACTTT 496  
Db 52178 AAAACAAAAAATAATCAAGGCTAGGTGGCGTGGCTCACACCTATAATCTCAGCACTTT 52119  
Qy 497 GGGAGGCTGAGCGAGGAGATCACCTTTGGTCAGGATTTTGGAGACTAGCCTGGGCCAACAT 556  
Db 52118 GGGAGGCAAGCGAGGAGATTAAGTGGTGGGAGTTTGGAGACCGCCTGGGCCAAT 52059  
Qy 557 GCGGAAACTCTGTCTTAAAAAAAATAACAAAAATAGCCTGGCATGGTGGCGGCGCCT 616  
Db 52058 GGTGAAACCCCATCTCT-ACTAAAAATACAAAAATTAACAGGTGTGGTGGTGCATGCCT 52000  
Qy 617 GTAATCTCAGTGTCTCAGGAGCGGAGACAAAGAAATCACTTAACCCGAGGTGGAGG 676  
Db 51999 GTAATCCAGCTGTCTCAGGAGGCTGAGGCAACAAAGAAATTTGCTTGAATCCAGGAGCAGAT 51940

Qy 677 TTCAGTGAGCTGAGATCGTGCCA-CTCACTCCAACCTGGGAGACAGAGTGACACTTTTG 735  
Db 51939 TTCTAGTGAGCCAAAGATCATGCGCACCACTCCAGCTGGGCAACAGACTGAGACTCTGT 51880  
Qy 736 TCTCAAAAGAAAAAACAACAA--GTAAAAAGAAACAGGTGAAGTTAACTTTAATAA 793  
Db 51879 CTCAAAAATAATAATAACTAATAAGGGGAAAAAAGAAACAGGTGAATAATTCATTGTAATAT 51820  
Qy 794 CCCAATGTATCCCAATACAAATCATTTCAAAGTGTAAATTAATAATAAACAATTAATGAATG 853  
Db 51819 GTCTTATTAACCGAATCTGTCTACACATTAATTTCAAACGTAATCGATCTAGGTTTATT 51760  
Qy 854 AGATACTTTACATCTCTTTCTGTTTTCATATTAAGTCTTTGAAAGTGAGTATATATGTT 913  
Db 51759 AAG-----ACAGTTTCCCTGTTTCTTACTAAGCTGGGAAAACTGGTATCGGTTCC 51706  
Qy 914 ATCTGACAGCACATCTCAATTTGGACTAGCTACATTTTCAAGTCTCAGTAGCCACATGT 973  
Db 51705 ACATCACAGTATGGCTCAGGGCGAACCCAGCCACATTTTCGGGTGCTTAGGGGCCACATAC 51646  
Qy 974 GGC 976  
Db 51645 TGC 51643

## RESULT 11

US-09-949-016-12893/c  
; Sequence 12893, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12893  
; LENGTH: 131724  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12893

Query Match 23.0%; Score 229.8; DB 4; Length 131724;  
Best Local Similarity 79.6%; Pred. No. 9.5e-50;  
Matches 284; Conservative 0; Mismatches 72; Indels 1; Gaps 1;  
Qy 436 CACATTAAAAACAGGTAAAAAGCTGGGCGAGTGGCTCACACCTGTAAATCCAGCACTT 495  
Db 5353 CAATTAATAAACACAACTACAGCGCGGCGAGTGGCTCACGCTGTAAATCCAGCACTC 5294  
Qy 496 TGGAGGCTGAGGAGGAGCATACCTTTGGTCAGGAGTTTGAGACTAGCCTGGCCACA 555  
Db 5293 TGGGAAGCCGAGGCGGTGGATCACCTGAGTCAAGGATTTGAGACCAGCCTGGCCATA 5234  
Qy 556 TGGCGAAACTCTGTCTTAAAAAATAACAAAAATTAGCCTGCGATGGTGGCGGCGCC 615  
Db 5233 TGGTGAATCCCGTCTCTAACTAAATAACAAAAATTAACCGGTGTGGTGGCGGTGTC 5174  
Qy 616 TGTAAATCTAGCTGCTCAGGAGCGGAGACACAAGAAATCACTTAAACCCAGAGGTGGAG 675  
Db 5173 TGTAGTCCAGCTACTCGGAGGCGGAGGAGGAAGAATCGCTTGAAACCCAGAGGCGAG 5114  
Qy 676 GTTGCACTGAGCTGAGATCGTCCACT-CACTCAAACCTGGGAGACAGAGTGACACTTTT 734  
Db 5113 GTTGCACTGAGCTGAGATCATGCCACTACAGCCTTGGGGACAGAGGAGACTCCT 5054



; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17446  
; LENGTH: 18319  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17446

Query Match 22.5%; Score 225.6; DB 4; Length 18319;  
Best Local Similarity 78.0%; Pred. No. 5.1e-49;  
Matches 284; Conservative 0; Mismatches 79; Indels 1; Gaps 1;  
Qy 416 ATTTCAGTATTTAGTACGACATTAAACAGGTAAAGGCTGGCGGAGTGGCTCA 475  
Db 11469 ATATCGAATATTAGCATTACGCTTTAAACGTGGGAGGAGACTGAGCGCGATGGCTCA 11528  
Qy 476 CACCTGTAATCCAGCACTTTGGGAGGCTGGGAGGAGGAGGCTGGCGGAGGTT 535  
Db 11529 CGCTGTAATCCAGCACTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGTT 11588  
Qy 536 TGAGACTAGGCTGGCCCAACATGGCGAACTGTGTCTCTAAATAAAATACAAAAATTAGC 595  
Db 11589 CGAGACCAGCTGACCAACATGGTGAACCTGTCTCTACTAAAAATACAAAGTTAGC 11648  
Qy 596 CTGGCATGGTGGCGGCGGCTGTGTAATCTCAGTCTCAGGAGGCGGAGACACAGAATCA 655  
Db 11649 TGGGCATGGTGGTGGATGCGCTGTAATCTCAGTCTCAGGAGGCTGAGGAGGAGATTG 11708  
Qy 656 CTTAAACCCAGGAGTGGAGGTTCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 714  
Db 11709 CTTGAACCCGAGGAGGAGGTTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 11768  
Qy 715 GGAGACAGAGTGACACTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAACAG 774  
Db 11769 TGTGACAGAATGAGACTGTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAAGGT 11828  
Qy 775 GTGA 778  
Db 11829 GGA 11832

RESULT 15  
US-09-768-185A-1  
; Sequence 1, Application US/09768185A  
; Patent No. 6818758  
; GENERAL INFORMATION:  
; APPLICANT: Cassel, Michael et al  
; TITLE OF INVENTION: Estrogen receptor beta variants and  
; TITLE OF INVENTION: methods of detection thereof  
; FILE REFERENCE: CL000280  
; CURRENT APPLICATION NUMBER: US/09/768,185A  
; CURRENT FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 09768185  
; PRIOR FILING DATE: 2001-01-24  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 325791  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-09-768-185A-1

Query Match 22.5%; Score 225.4; DB 4; Length 325791;  
Best Local Similarity 82.6%; Pred. No. 2e-48;

Matches 270; Conservative 0; Mismatches 56; Indels 1; Gaps 1;  
Qy 441 TAAACAGGTAAACAGGCTGGGCGGAGTGGCTCAGACCTGTATATCCAGCAGCTTTGGGA 500  
Db 282763 TGAAGAGAAACAGAGAGGCGGCGGAGTGGCTCAGACCTGTATATCCAGCAGCTTTGGGA 282822  
Qy 501 GGCTGAGGCGGAGGATCACCTTTGGTTCAGGAGTTTTCAGACTAGCCTGGCCCAACATGGCG 560  
Db 282823 GGCCAGGCGAGGTGGATCATCTGAGGTTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGCA 282882  
Qy 561 AAACCTGTCTCTTAAATAAAATACAAAAATTAGCCTGGCATGGTGGCGGCGCTGTAA 620  
Db 282883 AAACCTGTCTCTTAAATAAAATACAAAAATTAGCCTGGCATGGTGGCGGCGCTGTAA 282942  
Qy 621 TCTCAGCTGCTCAGGAGGCGGAGACACAGAATCACTTAAACCCAGGAGTGGAGGTTGC 680  
Db 282943 TCCAGCTGCTCGGGAGGCTGAAAGCAGGAGAAATCGCTTGAACCCAGGAGGAGGTTGC 283002  
Qy 681 AGTGAGCTGAGATCGTGCCACT-CAGTCCAAACCTGGGAGACAGAGTGCACACTTTTGTCTC 739  
Db 283003 AGTGAGCTGAGATCGCACCCTGCACCTCCAGCTTCCAGCTTGGGCGACAGATCGAGACTCTGTCTCA 283062  
Qy 740 AAAAAAGAAAAAACAAGTAAAAA 766  
Db 283063 AAAAAAGAAAAAACAAGTAAAAA 283089

Search completed: July 3, 2005, 18:09:02  
Job time : 117.609 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 400.727 Seconds

(without alignments)  
15667.668 Million cell updates/sec

Title: US-09-936-271c-13\_COPY\_6000\_7000

Perfect score: 1001

Sequence: 1 gacaacaaagagccccaag.....gttactgtatggatggcac 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*

13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*

14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*

15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*

16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*

17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*

18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*

19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*

20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*

21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*

22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*

23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*

24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*

25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252.6	25.2	52216	9	US-09-747-810-1
2	248.8	24.9	56098	13	US-10-087-192-1612
3	240.2	24.0	55827	9	US-09-813-133A-3
4	240.2	24.0	55827	14	US-10-212-877-3
5	240.2	24.0	55827	21	US-10-954-110-3
6	235.6	23.5	27189	13	US-10-087-192-1828
7	235.6	23.5	84105	19	US-10-741-601-5637

c	8	235.2	23.5	1980090	20	US-10-719-993-6815	Sequence 6815, Ap
c	9	235.2	23.5	1980090	21	US-10-741-600-17676	Sequence 17676, A
c	10	227.4	22.7	23071	9	US-09-764-864-1673	Sequence 1673, Ap
c	11	226.2	22.6	133955	13	US-10-087-192-1984	Sequence 1984, Ap
c	12	225.4	22.5	74424	21	US-10-773-678-153	Sequence 153, Ap
c	13	225.4	22.5	325791	11	US-09-768-185A-1	Sequence 1, Appli
c	14	225.2	22.5	35641	9	US-09-962-436-306	Sequence 306, Ap
c	15	225.2	22.5	35641	9	US-09-880-107-2225	Sequence 2225, Ap
c	16	225.2	22.5	35641	21	US-10-843-841A-2765	Sequence 2765, Ap
c	17	224.8	22.5	629	13	US-10-027-632-252422	Sequence 252422,
c	18	224.8	22.5	629	13	US-10-027-632-252423	Sequence 252423,
c	19	224.8	22.5	629	17	US-10-027-632-252422	Sequence 252422,
c	20	224.8	22.5	629	17	US-10-027-632-252423	Sequence 252423,
c	21	224.8	22.5	629	19	US-10-681-818-155	Sequence 155, Ap
c	22	224.8	22.5	629	19	US-10-681-818-156	Sequence 156, Ap
c	23	224.6	22.4	653122	13	US-10-087-192-226	Sequence 226, Ap
c	24	223.2	22.3	67253	22	US-10-737-082-88	Sequence 88, Appl
c	25	223.2	22.3	67253	22	US-10-765-790-88	Sequence 88, Appl
c	26	221.6	22.1	2591	19	US-10-755-889-83	Sequence 83, Appl
c	27	221.2	22.1	84409	19	US-10-741-601-5696	Sequence 5696, Ap
c	28	221.2	22.1	84409	21	US-10-741-600-17771	Sequence 17771, A
c	29	220.8	22.1	96587	11	US-09-997-722-250	Sequence 250, Ap
c	30	220.6	22.0	203264	13	US-10-087-192-988	Sequence 988, Ap
c	31	220.2	22.0	126990	19	US-10-717-597-68	Sequence 68, Appl
c	32	220	22.0	32146	9	US-09-764-860-797	Sequence 797, Ap
c	33	220	22.0	32146	14	US-10-074-095-797	Sequence 797, Ap
c	34	220	22.0	32146	17	US-10-212-872-797	Sequence 802, Ap
c	35	220	22.0	32248	9	US-09-764-860-802	Sequence 802, Ap
c	36	220	22.0	32248	14	US-10-074-095-802	Sequence 802, Ap
c	37	220	22.0	32248	17	US-10-212-872-802	Sequence 214, Ap
c	38	220	22.0	96589	18	US-10-052-482-214	Sequence 1470, Ap
c	39	219.2	21.9	100267	18	US-10-240-425-1470	Sequence 8960, Ap
c	40	218.6	21.8	4329	10	US-09-764-891-8960	Sequence 17953, A
c	41	218.6	21.8	43159	21	US-10-741-600-17953	Sequence 17953, A
c	42	218.6	21.8	106236	21	US-10-741-600-17759	Sequence 17759, A
c	43	218.6	21.8	366803	20	US-10-719-993-6805	Sequence 6805, Ap
c	44	218.6	21.8	410846	21	US-10-481-613-1	Sequence 1, Appli
c	45	218.4	21.8	2312	9	US-09-764-877-3606	Sequence 3606, Ap

ALIGNMENTS

RESULT 1

US-09-747-810-1

; Sequence 1, Application US/09747810

; Patent No. US20020012903A1

; GENERAL INFORMATION:

; APPLICANT: Goldman, Steven A.

; APPLICANT: Okano, Hideyuki

; TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGENITOR CELLS

; FILE REFERENCE: 19603/3580

; CURRENT APPLICATION NUMBER: US/09/747,810

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/173,003

; PRIOR FILING DATE: 1999-12-23

; NUMBER OF SEQ ID NOS: 2

; SEQ ID NO 1

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-747-810-1

Query Match 25.2%; Score 252.6; DB 9; Length 52216;

Best Local Similarity 69.1%; Pred. No. 9.8e-49;

Matches 403; Conservative 0; Mismatches 174; Indels 6; Gaps 4;

Qy	417	TTTCCAGTTTCTTAGTACCCACATTAAACAGGTAAACAGCTGGGCGCAGTGGCTCAC	476
Db	41515	TTAAATTTCTTAGGACACATTAA--NAGACATTAAGCCGGCGGTGGCTCAC	41572
Qy	477	ACCTGTAATCCAGCAGCTTTGGGAGGCTGAGGCGAGGACAGATCACCTTTGGTGAGGAGTTT	536











Db 538803 ATAGTCTATGTTCTTTTTCCTAAG-----TCTGAAATTCGGTGTGTAT-TTAC 538756  
Qy 916 GCTGACAGCATCTCAATTGGACTAGTACATTTTCAGGTGCTCAGTAGCCATGCG 975  
Db 538755 ACTGACAGCAUCTCAATTAGACTAGCCACATTTCCAGTGTGCTGCTGATGTAC 538696  
Qy 976 CTAGCAGTTACTGTATTGGATGGCAC 1001  
Db 538695 CTGTGGCTACCCACATGGACAGCAC 538670

RESULT 10  
US-09-764-864-1673/c  
; Sequence 1673, Application US/09764864  
; Patent No. US20020132753A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT223  
; CURRENT APPLICATION NUMBER: US/09/764,864  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1792  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1673  
; LENGTH: 23071  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-864-1673

Query Match 22.7%; Score 227.4; DB 9; Length 23071;  
Best Local Similarity 70.1%; Pred. No. 6.3e-43;  
Matches 445; Conservative 0; Mismatches 156; Indels 34; Gaps 9;

Qy 381 GTGTTTAGTGAATAATAGCTGAGCTGTTATGTCATTTCCAGTTTTTTAGTAGCCACAT 440  
Db 10748 GTGCTAAGAGAAATATAATGTGAGCCACATATATAATTTTAAATTTCTCGCTCCACATT 10689  
Qy 441 TAAACAG-GTAAAAAGGCTGGCGAGTGGCTCACACCTGTAAATCCCGACATTTGGG 499  
Db 10688 AACAAAAGTGAAGAGGCTGGCGCGGTGCTGCTGCTGTAAATCCCGACATTTGGG 10629  
Qy 500 AGGCTGAGGCGAGCATCACTTTTGTGTAGGAGTTTGTAGACTAGCCCTGGCCACATGGC 559  
Db 10628 AGGCTGAGAGCGGATCACTGAGTGGAGTTCGAGACTAGCCTAACCAACATGGA 10569  
Qy 560 GAAACTCTGTCT---CTAATAAAATATACAAAATTTAGCCTGGCATGGTGGCGGCGCCT 616  
Db 10568 GAAATCCCATCTCTACTATAAAAAAAATTTACAAAATTTAGCTGGCGGTGGTGGCATGCCT 10509  
Qy 617 GTATCTCAGCTGCTCAGGAGCGGAGACACAAGATCACTTTAAACCCAGAGGTGGAGG 676  
Db 10508 GTATCCAGCTACTAGGAGGCTGAGGAGAGATCGCTTTGAACCCGGAGGCGGAGG 10449  
Qy 677 TTGCAGTGAGCTGAGATCGTGCCACT-CACCTCCAACTGGGAGACAGAGTGACACTTTTG 735  
Db 10448 TTGCGGTGAGCGGAGATTGCGCCATTGCATCCAGCTGGCCACAGAGCGAACTCCA 10389  
Qy 736 TCTCAAAAGAAAAAACAAGTAA-----AAAAAACAAGTGAAGTTAACTTT 788  
Db 10388 TCTCAAAAAAAACAAAAAACAACAAAAAAGTGAAGGAAACAGGGAAGTTAAATTTA 10329  
Qy 789 AA-----TAACCCATGTATCCCAATACATCAATTTCAAGTCTAAATATATAAACA 843  
Db 10328 AATAGTTTAAACCAAGTGTATCCATGATATATTTTCAACATGTATTAATCAACATTA 10271  
Qy 844 ATTATGATGAGATCACTTTTACATTTCTTTCTTTTGTATTTTCAATTAAGTCTTTTGAAGTGAG 903  
Db 10270 ACTATTACTGAGACATTTTACATTTGTT-----TATACATCTCTCAAAATCCAT 10223  
Qy 904 TATATATGTTATGCTG-ACAGCAGATCTCAATTTGGAGTAGCTACATTTAGGTGCTCAG 962  
Db 10222 TGTGTATTTTCTACTTTCACAGCAACATCAATTTGGG--CAGCCACATTTCAAAATTTCTCGA 10165

Qy 963 TAGCCACATGTGGCTAGCAGTTACTGTATTGGATG 997  
Db 10164 TACCACAAATGCTATTGGCTATTATTGTTTG 10130

RESULT 11  
US-10-087-192-1984  
; Sequence 1984, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1984  
; LENGTH: 133955  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-1984

Query Match 22.6%; Score 226.2; DB 13; Length 133955;  
Best Local Similarity 68.4%; Pred. No. 2.6e-42;  
Matches 357; Conservative 0; Mismatches 158; Indels 7; Gaps 3;

Qy 425 TTTTGTAGTAGCCACATTAACACAGGTAA-----AAAAGGCTGGCGCAGTGGCTCACACC 479  
Db 115167 TTTCTAGTGGCCACATTTTAAAAACATAAGAAGAACAGGCGCGTGGTGCATGCC 115226  
Qy 480 TGTAAATCCAGCACTTTGGGAGGCTGAGCGAGGAGATCACTTTGGTCAGGAGTTTGAG 539  
Db 115227 TGTAAATCCAGCACTTTGGGAGGCTGAGGCAAGCAATCACTTTGAGGCCAGGAGTTCAAG 115286  
Qy 540 ACTAGCTGGCCCAACATGGCGAAAACCTGTCTCTTAAAAAAAATAACAAAAATTAGCCTGG 599  
Db 115287 ACCAGCCTGTCAACATGGTGAAACTCTGTCTCT-CTAATAAATAACAAAAATTAGCCAGG 115345  
Qy 600 CATGTGGCGGGCGCCTGTGTAATCTCAGCTGCTCAGGAGGCGGAGACACAGAAATCACTTA 659  
Db 115346 CATGGTGGCGCACACCTGTAAATCCAGTACTCAAGTAGGTGAGGCAACAGAAATCACTTG 115405  
Qy 660 AACCCAGGAGGTGGAGGTTGCAGTGTGAGTGTGAGTCTGAGATCGTGCCACT-CACTCCAACTGGGAG 718  
Db 115406 AACCCGGAAGCGAGGTTGCGAGTGAGCCGAGATCACGCCACTGTACTTCCAGCCTGTGTG 115465  
Qy 719 ACAGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAAGAAAAACAGGTGA 778  
Db 115466 ACAGAGCAAGACTCTGTCTCAAAAAGAAAAAACAAGTAAAAAAAGAAAAACAATGGA 115525  
Qy 779 AGTTAACTTTAATAACCCAAATGTATCCCAATACATCAATTTCAAGTGTAAATATA 838  
Db 115526 ATTATTTTAACTAATCAATTTTATTTAAACCCAGTACTCAAAAGATTATCATTTTCAAAATG 115585  
Qy 839 AAACAATTTAATGATGAGATCTTTTACATCTCTTTTGTGTTTTCATATTTAAGTCTTTGAAA 898  
Db 115586 CAATTCATATATATTAATTTAGATATTTTACATTTCTCTTCTTAATGCTATGCTTTGACA 115645  
Qy 899 GTGAGTATATATGTTATGCTGACAGCAGCATCTCAATTTGGAC 940  
Db 115646 TCATGTGTGTGTTTATACATTTACAGCCCAATTTTAAATTCGGAC 115687

RESULT 12  
US-10-773-678-153/c

; Sequence 153, Application US/10773678  
; Publication No. US20050074879A1  
; GENERAL INFORMATION:  
; APPLICANT: Karras, James G  
; TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3  
; FILE OF INVENTION: Expression  
; FILE REFERENCE: ISPH-0828  
; CURRENT APPLICATION NUMBER: US/10/773,678  
; CURRENT FILING DATE: 2004-02-06  
; PRIOR APPLICATION NUMBER: 10/713,139  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: 09/758,881  
; PRIOR FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: PCT/US00/09054  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: 09/288,461  
; PRIOR FILING DATE: 1999-04-08  
; NUMBER OF SEQ ID NOS: 402  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 153  
; LENGTH: 74424  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; OTHER INFORMATION: Antisense oligonucleotide  
US-10-773-678-153

Query Match 22.5%; Score 225.4; DB 21; Length 74424;  
Best Local Similarity 68.3%; Pred. No. 3.1e-42;  
Matches 372; Conservative 0; Mismatches 166; Indels 7; Gaps 4;  
  
Qy 340 CAGGACCCCTCTTCCACAGAGCTATACCCCTAGAGTAGTGTGTAGTGAATAATG 399  
Db 38162 CAGGAGTCTTTACATTTGAATTAATACTAAGGCGAGCATCCCAACAGAAATGAA 38103  
  
Qy 400 CTGAGCTGCTTATGTCTATTTCCAGTTTCTTGTAGTACCACTTAAACAGGTAAAAAGGC 459  
Db 38102 TACAGGCCACACATGCTATTTAAATTTCTGTAGCCATCTTAAAGTTTAAAGAGGC 38043  
  
Qy 460 TGGGCGCAGTGGCTCACACCTGTATATCCAGCATTTTGGGAGGCTGAGGCGAGGATCA 519  
Db 38042 TGGGTGCGAGTGGCTCAGCGCTGTAATCCAGAACTTTGGGAGGCCAAGGCGAGGATCA 37983  
  
Qy 520 CTTTGTGTCAGAGTTTGAGACTAGCTGGCCCAATGCGCAACTGTCTCTAAAAAA 579  
Db 37982 --TGAGGTGAGGATCGAGACCATCTCTGGCCAAATATGTTGAAACCCCGTCTCT-ACTAA 37926  
  
Qy 580 AAATACAAAAATTAGCCTGGCATGTGTGGCGGCGCTGTATCTCAGCTGCTCAGGAGGC 639  
Db 37925 AAATACAAAAATTAGCTGGTGTGGTGCCATGCTTTTATCCAGCTACTAGGAGGC 37866  
  
Qy 640 CGAGACACAAGATCACTTAAACCCAGGAGGTGGAGGTTCAGTGAGCTGAGATCGTGCC 699  
Db 37865 TGAGCGCAAGAATCGCTTGAACCCAGGAGGAGGTTGCACTGAGCGAGATCGTGCC 37806  
  
Qy 700 ACTCACTCAACCTGGGAGACAGAGTGACACTTTTGTCTCAAAAAGAAAAAAACAA 759  
Db 37805 ACTCACTCCAGCTGTGATGACATAGCAGAGAC-TCCATCTCAAAAAGAAAAAAAGAA 37747  
  
Qy 760 GTAAAAAGAACAGGTGAAGTTAACTTTAATAACCCCAATCTATCCCAATACATCAT 819  
Db 37746 AAAGAAAGAACAGGTGAAGTTAATTTAAT---GTATTTTATTTCAATTAATATATCC 37690  
  
Qy 820 TCAAAGTGAATTAATAAAAAAATATGAATGAGATACCTTACATCTCTTTCTTTT 879  
Db 37689 AAATGTTATTGCAACATATGACATACAAACATATTAATTAATTAATTAATTTT 37630  
  
Qy 880 TCATA 884  
Db 37629 GGAGA 37625

RESULT 13

Query Match

22.5%; Score 225.2; DB 9; Length 35641;

US-09-768-185A-1  
; Sequence 1, Application US/09768185A  
; Publication No. US20040185439A1  
; GENERAL INFORMATION:  
; APPLICANT: Cassel, Michael et al  
; TITLE OF INVENTION: Estrogen receptor beta variants and  
; FILE OF INVENTION: methods of detection thereof  
; FILE REFERENCE: CL000280  
; CURRENT APPLICATION NUMBER: US/09/768,185A  
; CURRENT FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 09768185  
; PRIOR FILING DATE: 2001-01-24  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 325791  
; TYPE: DNA  
; ORGANISM: HUMAN  
US-09-768-185A-1

Query Match 22.5%; Score 225.4; DB 11; Length 325791;  
Best Local Similarity 82.6%; Pred. No. 5.8e-42;  
Matches 270; Conservative 0; Mismatches 56; Indels 1; Gaps 1;  
  
Qy 441 TAAACAGGTAAAAAGCTGGCGCAGTGGCTCACACCTGTATATCCAGCACTTTGGGA 500  
Db 282763 TGAAGAAGAAAAAGAGAGCGCGCGCAGTGGCTGTATATCCAGCACTTTGGGA 282822  
  
Qy 501 GGCTGAGCGAGGAGAGATCACCTTTGGTCAGGAGTTTGAGACTAGCTGGCCCAACATGGCG 560  
Db 282823 GGCCAGCGAGGTGATCATCTGAGGTTCAGAGTTCAAGACTAGCTGGCCCAACATGGCA 282882  
  
Qy 561 AAATCTGTCTTAAAAAATAACAAAAATAGCTTGGCATGTGGCGGCGCTGTAA 620  
Db 282883 AAATCTGTCTTAAAAAATAACAAAAATAGCGAGGATGTGGCGGCGCTGTAA 282942  
  
Qy 621 TCTCAGCTGCTCAGGAGCGGAGACACAGAAATCACTTAAACCCAGGAGTGGAGGTTC 680  
Db 282943 TCCAGCTGCTCGGAGGCTGAAGCAGAGAAATCGCTTGAACCCAGGAGGAGGTTC 283002  
  
Qy 681 AGTGAGCTGAGATCGTGCCACT-CACCTCAACCTGGGAGACAGAGTGCACACTTTGTCTC 739  
Db 283003 AGTGAGCTGAGATCGCACCACTGCACCTCCAGCTTGGCGCAGATCGAGACTCTGTCTCA 283062  
  
Qy 740 AAAAAGAAAAAACAAGTAAAAA 766  
Db 283063 AAAAAGAAAAAACAAGTAAAAA 283089

RESULT 14

US-09-962-436-306  
; Sequence 306, Application US/09962436  
; Patent No. US20020081301A1  
; GENERAL INFORMATION:  
; APPLICANT: Soppet, Daniel  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE OF INVENTION: Sets  
; FILE REFERENCE: 689290-75  
; CURRENT APPLICATION NUMBER: US/09/962,436  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,082  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/234,924  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 568  
; SOFTWARE: Patent In version 3.0  
; SEQ ID NO 306  
; LENGTH: 35641  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-962-436-306

Best Local Similarity	84.4%;	Pred. No. 2.5e-42;			
Matches	265;	Conservative	0;	Mismatches	48;
				Indels	1;
				Gaps	1;
QY	455	AAGGCTGGGGCGCAGTGGCTCACACCTGTATCCAGCACCTTTGGGAGGCTGGAGCGAGGCA	514		
Db	14689	AAGGCTGGGCTCAGTGGCTCACGCTGTATCCAGCACCTTTGGGAGGCTGGAGCGAGGCC	14748		
QY	515	GATCACCTTTGGTCAGAGAGTTTGAGACTAGCTGGCGCAATGCGGAAACTCTGTCTCTTA	574		
Db	14749	GATCACCTGAGGTGAGGAGTTTGAGACCAAGCCCTGGCCAAATATGTTGAAACCCCGTCTCTTA	14808		
QY	575	AAAAAAATACAAAAATTAGCCTGGCATGTGTGGCGGCGCTGTAATCTCAGCTGCTCAG	634		
Db	14809	TTAAAAATACAAAAATTAGCCAGGCGGTGTGGCGGCGCCTGTATCCAGCTACTCAG	14868		
QY	635	GAGCGCGAGACACAAGAAATCACTTAAACCCAGGAGGTGGAGGTTGCAGTGCAGCTGAGATC	694		
Db	14869	GAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGGAGCGGAGGTTGCAGTGCAGCTGAGATC	14928		
QY	695	GTGCCACT - CACTTCCACCTGGGAGACAGAGTGCACACTTTTGTCTCAAAAAAGAAAAAAA	753		
Db	14929	ACACTACTGCATCTCCAGCCTTGGGCAACAGAGTGAGACTCGGTCTCAAAAAAGAAAAAAA	14988		
QY	754	AAACAAAGTAAAAAA	767		
Db	14989	AAAAAACACAAAAA	15002		

Db

14869

GAGGCTGAGGCGAGGAGAAATCACTTGAACCCAGGAGCGGAGGTTGCAGTGCAGCTGAGATC

14928

QY

695

GTGCCACT - CACTTCCACCTGGGAGACAGAGTGCACACTTTTGTCTCAAAAAAGAAAAAAA

753

Db

14929

ACACTACTGCATCTCCAGCCTTGGGCAACAGAGTGAGACTCGGTCTCAAAAAAGAAAAAAA

14988

QY

754

AAACAAAGTAAAAAA

767

Db

14989

AAAAAACACAAAAA

15002

Search completed:

July 4, 2005, 02:13:01

Job time :

410.727 secs

Search completed: July 4, 2005, 02:13:01  
Job time : 410.727 secs

RESULT 15

```

US-09-880-107-2225
; Sequence 2225, Application US/09880107
; Patent NO. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2225
; LENGTH: 35641
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 176568
US-09-880-107-2225

```

	Query Match	22.5%;	Score	225.2;	DB	9;	Length	35641;
	Best Local Similarity	84.4%;	Pred. No.	2.5e-42;				
	Matches	265;	Conservative	0;	Mismatches	48;	Indels	1;
Qy	455	AAGGCTGGGCGCAGTGGCTCACACCTGTGTAATCCAGCACTTTGGGAGGCTGAGGCAGGCA	514					
Db	14689	AAGGCTGGGCTCAGTGGCTCAGCTCTGTAATCCAGCACTTTGGGAGGCTGAGGCAGGCC	14748					
Qy	515	GATCACCTTGGTCAGGAGTTTGAGACTAGCTGGCCACATGCGGAACCTCTGCTCTCA	574					
Db	14749	GATCACCTGGGTCAGAGTTTGAGACCACTGGCCCAATATGGTGAACCCCGTCTCTA	14808					
Qy	575	AAAAAAATACAAAAATTAGCTGGGCATGGTGGCGGCGCCCTGTGTAATCTCAGCTGCTCAG	634					
Db	14809	TTAAAAATACAAAAAATTAGCCAGGCGGTGGTGGCGGCACTGTGTAATCCAGCTACTCAG	14868					
Qy	635	GAGGCGCGAGACACAAGAATCACCTTAAACCCAGAGGTGGAGGTTGCGAGTGAGCTCAGATC	694					

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 2176.56 Seconds  
(without alignments)  
17505.766 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_6000\_7000

Perfect score: 1001

Sequence: 1 gacaaccaagagcccccag.....gttactgtatggatggcac 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_est4:\*

5: gb\_est5:\*

6: gb\_est6:\*

7: gb\_est7:\*

8: gb\_est8:\*

9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	226.8	22.7	908	6	CD245375 AGENCOURT
2	223	22.3	694	8	BZ603010 WHAAX35TF
3	221.6	22.1	1750	3	AF119908 Homo sapi
4	220	22.0	769	6	CD370173 UI-H-Ftl1
5	219.6	21.9	500	5	BX491983 DKFZp781O
6	218.4	21.8	837	2	BE790769 601581950
7	216	21.6	649	8	B71494 RPC111-8N9
8	216	21.6	1019	6	CD518470 AGENCOURT
9	215.6	21.5	2449	3	CR749224 Homo sapi
10	215.4	21.5	573	5	BX477553 DKFZp686H
11	215.4	21.5	823	6	CD101954 AGENCOURT
12	215	21.5	397	1	AV759632 AV759632
13	215	21.5	587	5	EX644875 DKFZp781D
14	215	21.5	633	7	CK905714 1j76h09.x
15	214	21.4	4991	2	AW516097 Homo sapi
16	214	21.4	4593	3	HS8803561
17	213.4	21.3	517	8	AQ264959 CITBI-EI-
18	213	21.3	569	8	AQ490878 RPC111-2
19	212.8	21.3	700	5	EX508919 DKFZp686G
20	212.6	21.2	790	6	CD643827 AGENCOURT
21	212.4	21.2	619	5	EX485408 DKFZp686B
22	211.8	21.2	576	6	CD652305 AGENCOURT
23	211.8	21.2	731	7	CN307840 170004182
24	211.8	21.2	15970	8	AQ839852 260U13-C5

C 25	211.4	21.1	537	8	AQ390925	AQ390925 CITBI-EI-
26	211.4	21.1	624	9	AG019812	AG019812 Homo sapi
27	211.4	21.1	654	7	CK003802	CK003802 AGENCOURT
28	211.4	21.1	687	6	CD237038	CD237038 FNPAPH09
29	211.4	21.1	692	5	BU662309	BU662309 C184C05.2
C 30	211.2	21.1	338	2	AW502796	AW502796 UI-HF-BR0
C 31	211.2	21.1	596	8	AQ320557	AQ320557 RPC111-96
32	211.2	21.1	801	8	BZ600703	BZ600703 WHASC76TF
33	211	21.1	552	8	B68316	B68316 CIT978SK-A-
C 34	210.8	21.1	657	7	CF123536	CF123536 UI-HF-CHO
C 35	210.6	21.0	1016	5	BQ674146	BQ674146 AGENCOURT
C 36	210.4	21.0	388	2	AW069227	AW069227 cr41h09.x
C 37	210.4	21.0	735	7	CF146929	CF146929 UI-HF-CB0
C 38	210.2	21.0	671	9	AG079103	AG079103 Pan trogl
C 39	210.2	21.0	849	6	CB991080	CB991080 AGENCOURT
C 40	209.8	21.0	397	1	AI421950	AI421950 wv1d10.x
C 41	209.8	21.0	463	2	AW026305	AW026305 wv1d10.x
C 42	209.6	20.9	461	8	B65075	B65075 CIT-HSP-201
C 43	209.6	20.9	501	5	EX485916	EX485916 DKFZp686E
C 44	209.6	20.9	604	5	BM990713	BM990713 UI-H-DIO-
C 45	209.2	20.9	534	6	CB050438	CB050438 NISC_gj17

ALIGNMENTS

RESULT 1  
CD245375/c  
LOCUS CD245375 908 bp mRNA linear EST 22-MAY-2003  
DEFINITION AGENCOURT\_14128276 NIH\_MGC\_181 Homo sapiens cDNA clone  
IMAGE:30374350 5', mRNA sequence.  
ACCESSION CD245375  
VERSION EST.  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov  
Plate: NDAM433 row: j column: 23  
High quality sequence start: 10  
High quality sequence stop: 557.  
Location/Qualifiers  
1. 908  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="IMAGE:30374350"  
/tissue\_type="White Matter"  
/dev\_stage="Unknown"  
/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances) "  
/note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 Kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 22.7%; Score 226.8; DB 6; Length 908;  
Best Local Similarity 69.3%; Pred. No. 4.6e-25;



	Matches	384;	Conservative	0;	Mismatches	157;	Indels	13;	Gaps	5;
Qy	437	ACATTAAAA	CAGGTAAAA	AGGCTGGG	CGCAGTGGCT	CACACCTGT	TAATCC	CACACATTT	496	
Db	680	AAATGTAA	AGAAAAAGT	CGGCCGG	CGCGTGGCT	CAGCCTGT	TAATCC	CACACATTT	621	
Qy	497	GGGAGGCT	GAGCGCAGAT	CACCTTT	TGTCAGGAGTTT	TGCAGACTAGC	CTTGCCCA	CAT	556	
Db	620	GGGAGAC	CGGAGCGAGG	CGGATC	AC--GACGT	CGGAATCGAG	ACCATCTCTT	GTTTAA	CAC	563
Qy	557	GGCGAA	ACTCTGCTCT	TAATAAAAA	ATAAAAAAT	TAGCCTTGG	CATGGTGGCGG	CGCCT	611	
Db	562	GGTGA	AACTCGTCTCT	TACTTAA	AAATACAAAA	AAATTAGC	CAGGCATGGTGGCG	ATGCCT	503	
Qy	617	GTAATCT	CAGCTGCTCAGG	AGGCGCAGAC	ACAAGAAT	CACTTAA	ACCACGAGGATGG	AGG	676	
Db	502	GTAGTCC	CAGCTACTTGG	AGGCTGAGC	CAGGAGAT	TGGCTGA	ACCCGGGAGG	CAGAC	443	
Qy	677	TTCCAGT	GAGCTGAGAT	CGTGCC	ACTCCTCA	ACCTGGG	AGACAGTGN	ACATTTTGT	736	
Db	442	TTGCAGT	GTAGCTGAG	ATAGC	ACCCTGC	ACTCCAC	CGCTGGGCG	CACAGAC	384	
Qy	737	CTCAAA	AGAAAA	AAAAAACA	AGTAAAA	AGAAACAGG	TGAAGTTAA	CTTTAATA	796	
Db	383	CTGTCT	CAGAAAA	AAAAAAGAAAA	AAAAAATA	AGTCAT	AGATTTTA	TATATT	324	
Qy	797	AATGTAT	CCCAATAC	AAATCA	TTCATTT	CAAGTGA	ATTAATA	TAAACCAAT	856	
Db	323	AGTATGT	CTAAAAAT	TATCAT	TTAA-----	CATGTA	ATCAACAA	AGTTATT	270	
Qy	857	TACTTT	ACATTTCTTT	CTGTGTTT	CATATTA	AGCTTTT	GAAAGTGAGT	ATATATGT	915	
Db	269	TATTCT	ACATTCCT	GTGTGG--	GGGTA	ATAAATCTTT	GAAATTA	ATGTGAT	213	
Qy	916	GCTGAC	AGCACATCTCA	ATTGG	ACTAGCTAC	ATTTCAGG	TGCTCAG	TAGCCACAT	975	
Db	212	ACTTAC	AGCTGATCTT	GATTCAG	ACTAG	CCACATTTCC	AGTGTCTCAG	TAGCCCCC	153	
Qy	976	CTAGC	AGTTACTGT						989	
Db	152	CTAGC	AACTATGTT						139	

RESULT 2	BZ603010	694 bp	DNA	linear	GSS 08-JUN-2003
LOCUS	WHAAX35TF	Human MCF7 breast cancer cell line library (MCF7_1)	Homo sapiens genomic clone MCF7_1-6F22	genomic survey sequence.	
DEFINITION					
ACCESSION	BZ603010				
VERSION	BZ603010.1	GI:31511472			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 694)				
AUTHORS	Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C.				
TITLE	End-sequence profiling: Sequence-based analysis of aberrant genomes				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)				
MEDLINE	22709111				
PUBMED	12788976				
COMMENT	Contact: Volik SV Colin Collins' lab UCSF Comprehensive Cancer Center UCSF Box 0808, San Francisco, CA 94143-0808, USA Tel: 415 502 7066 Fax: 415 502 5685 Email: svolik@cc.ucsf.edu This clone is available from Amplicon Express <a href="http://www.genomex.com">http://www.genomex.com</a> Class: BAC ends.				

FEATURES  
 source  
 Location/Qualifiers  
 1. .694  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="MCF7 1-6F22"  
 /sex="female"  
 /clone\_lib="Human MCF7 breast cancer cell line library (MCF7 1)"  
 /note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN  
 Query Match 22.3%; Score 223; DB 8; Length 694;  
 Best Local Similarity 79.1%; Pred.No. 1.9e-24;  
 Matches 265; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
 QY 439 ATTAAACAGGTAATAAAGGCTGGGCGCAGTGCTGCCTACACCTGTATATCCAGCACATTTGG 498  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 28 ACTAAATGGAATGTTGCCITGGGCGGCTGCTCAGGCTGTATATCCAGCACATTTGG 87  
 QY 499 GAGGCTGAGCGAGCAGATCACTTTGGTCAGAGATTGAGATAGCTCTGGCCAAACATGG 558  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 88 GAGGCTGAGTTGGGCAGATTGCTGAGGTTCAGAGTTCAGACAGCCTGGGCAACATGG 147  
 QY 559 CGAAACTCTGTCTCTAAATAAATAACAAATATTAGCTGGCATGTTGGCGGCGCCTGT 618  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 148 CGAAACCCCTGTCTCTACTAAATAATACAAATAATTAGCTGGCATGTTGGCGGCTGCTGT 207  
 QY 619 AATCTCAGCTGTCTCAGGAGGCGCAGACACAAAGAAATCACTTAAACCCAGAGGTGGAGGTT 678  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 208 AATCAGCTACTTCGGAGGCTGAGGCAAGAGAAATCGTTGAGCCAGAGGCGAGAGTT 267  
 QY 679 GCAGTGAGCTGAGATCGTGCCACTACTCAACTGGGAGACAGATGACACTTTTGTTCT 738  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 268 GCAGTGAGCGAGATCGCGCACTGCCCTCAGGCTGGGTGACAGAGTAAGACTCTGTCTC 327  
 QY 739 CAAAAGAAAAAATAAACAAGTAAAAAGAAACA 773  
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
 328 AAAAAAATAAATAAATAAATAAATAAATAAATAA 362

RESULT 3  
 AFL19908  
 LOCUS Homo sapiens PRO2955 mRNA, complete cds.  
 DEFINITION  
 ACCESSION AFL19908  
 VERSION AFL19908.1 GI:7770252  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.  
 1 (bases 1 to 1750)  
 TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver  
 JOURNAL Unpublished  
 REFERENCE  
 AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.  
 2 (bases 1 to 1750)  
 TITLE Direct Submission  
 JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China  
 FEATURES  
 source  
 Location/Qualifiers  
 1. .1750  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"



```
Db 626 CTCAAAAAAGAAAAAGATGAGAAATAATCAA 661

RESULT 5
BX491983
LOCUS
DEFINITION DXFzp78100513_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
VERSION DXFzp78100513 5', mRNA sequence.
BX491983
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Fodor, G., Han, M. and Wiemann, S.
TITLE EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
COMMENT Unpublished (2003)
CONTACT: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Heinrich-
Heine-University, Duesseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No sl sequence available.
This clone (DKFzp78100513) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFzp78100513"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"

ORIGIN
Query Match 21.9%; Score 219.6; DB 5; Length 500;
Best Local Similarity 82.7%; Pred. No. 7e-24; Indels 1; Gaps 1;
Matches 263; Conservative 0; Mismatches 54;

Qy 443 AACACAGGTAAAGAGGCTGGCGCAGTGGCTTCACACCTGTAAATCCAGCACTTTGGGAGG 502
Db 13 ATAAAGGAATAGAGGCTGGCGCATGGTGGCTCACGCGTGTAAATCCAGCACTTTGGGATG 72
Qy 503 CTGAGGAGGCGAGATCACTTTGGTTCAGGAGTGTGAGACTAGCTGGCGCAACATGGCGAA 562
Db 73 CCGAGGGCGGCGAGATCACTGAGGTTCAGGCTTCGAGACTAGCTGGCGCAACATGGTGAA 132
Qy 563 ACTGTGCTCTAAAAAATAACAAAAATACAAAAATAGCTGGCATGTGGCGGCGCGCTGTAATC 622
Db 133 ACCCCATCTTTACTAAAAAATCAAAAAATAGCCGGGTGTGGTGGAGACCCCTGTAAATC 192
Qy 623 TCAGCTGCTCAGGAGGCGGAGACACAAAGAAATCACTTAAACCCAGAGGTGGAGGTTCGAG 682
Db 193 CCAGCTACTTGGGAGGCTGAGAGGAGAGATCGCTTGAACGTAGAGGTGGAGGTTCGAG 252
Qy 683 TGAGCTGAGATCGTGCCACT-CATCCAACTGGGAGACAGAGTGACACTTTTGTCTCAA 741
Db 253 TGAGCTGAGATGTGCCACTGCACTAGCTAGCTGGGTGAGAGTGACAGAGTGAGACTCTTGTCTCAA 312
Qy 742 AAGAAAAAATAACAA 759
Db 313 AAAAAAATAAAAAA 330
```

---

```
BE790769 837 bp mRNA linear EST 20-SEP-2000
601581950F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936325 5',
mRNA sequence.
BE790769
VERSION BE790769.1 GI:10211967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM780 row: c column: 14
High quality sequence stop: 723.

FEATURES
Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936325"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

---

```
Query Match 21.8%; Score 218.4; DB 2; Length 837;
Best Local Similarity 70.0%; Pred. No. 8.9e-24;
Matches 308; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

Qy 332 AAGACAGTCAGGACCCCTGTTCTCACAGAGCTCATACCTAGTAGTGGTGTAGTAG 391
Db 472 AAAAAAGTGAAGACCCCTGTTATGCGGTGAATCATGTCTGCCTGAATTCCTATGTTGAAG 413
Qy 392 AAATAAATGCTGAGTGTCTATGTCATTTCCAGTTTTTTAGTAGCCACATTAACACAGTA 451
Db 412 CCCTAACCCAGTAGCTCAGATGTGACTGTATTCATATAGGGCTTTAAGGGGTA 353
Qy 452 AAAAAAGCTGGCGCAGTGGCTCACCTGTGTATCCAGCACTTTGGAGGCTGAGGCAG 511
Db 352 ATTAAGGCTGGTGCAGTGGCTCACACCTATATCCCATCACTTTGGAGGCCGAGGCAG 293
Qy 512 GCAGATCACTTTGGTCAGGAGTTGAGCTAGCTGGCCCAACATGGGAACTCTGCT 571
Db 292 GTGTATCACTTGGAGTGGAGTTTGAGACAGCTGACCAACATGAAGAACCTGTCT 233
Qy 572 CTAATAAAAAATAACAAAAATTTAGCCTGGCATGTGGCGGGCGCTGTAAATCTCAGCTGCT 631
Db 232 CTACTAAAAATAACAAAAATTTAGCTGTGTGGTGTACAGCCCTGTAAATTCAGCCACT 173
Qy 632 CAGGAGCCGAGACACAAGAAATCACTTAACCCAGGAGGTGGAGTTGCACTGAGCTGAG 691
```

Db 172 TGGGAGGCTAAGGCTGGAGATCACCTGAAACCCGGGAGACAGAGTTCCAGTGAGCTGAG 113

Qy 692 ATCGTGCCACT-CACTCCAACTGGGAGACAGAGTGACACTTTTGTCTCAAAAAAGAAAA 750

Db 112 ATCGCGCACTGCACCTCAAGCTGGTGACAGAGCGAGACTCTGTCTCAAAAAACRAAG 53

Qy 751 AAAAAACAAGTAAAAAGAA 770

Db 52 AAAAAAAGAGGTAATTA 33

RESULT 7  
B71494/c

LOCUS B71494

DEFINITION RPC111-8N9.TP RPC11-11 Homo sapiens genomic clone RPC11-11-8N9, genomic survey sequence.

ACCESSION B71494

VERSION B71494.1 GI:2710715

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 649)

AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.

TITLE Use of BAC End Sequences for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Other GSSs: RPC111-8N9.TV  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are derived from the human BAC library RPC11-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1. .649  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7503008"  
/db\_xref="taxon:9606"  
/clones="RPC11-11-8N9"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPC11-11"  
/note="Vector: pBACes3.6; Site 1: EcoRI; Site 2: EcoRI; RPC11 Human Male BAC Library"

ORIGIN  
Query Match 21.6%; Score 216; DB 8; Length 649;  
Best Local Similarity 78.2%; Pred. No. 2.3e-23;  
Matches 297; Conservative 0; Mismatches 80; Indels 3; Gaps 3;

Qy 412 TGTCTATTCACAGTTTCTTTAGTAGCCACATTAAACAGGTAAAGGCTGGCGCAGTGG 471

Db 553 TGTACTATACGAATATATGAGGAGCTTTAAAGAGCTTTTACAGCGCGGTGTGGTG 494

Qy 472 CTCACACTGTAATCCAGCACTTTGGAGGCTGAGCGAGCATCACCTTTGGTCAGG 531

Db 493 CTTACACTGTAATCCAGCACTTTGGAGGCTGAGCGAGCATCACTTTGAGGTGAG 434

Qy 532 AGTTTGAGCTAGCTGGCCACATGGCGAACTCTGTCTCAAAAAAATACAAAAAT 591

Db 433 AGTTGGAGACAGCCTGGCCAAATGCGCAAAATCCTGTCTCT-ACTAAAAATACAAAAAT 375

Qy 592 TAGCCTGCATGTGTGGCGGGCGCCTGTAAATCTCAGCTGCTCAGGAGCCGAGACACAAGA 651

Db 374 TAGCCAGAGTGGTGGCGCATGCCTATAATCCAGCTACTTTAGGAGGCTGAGGCGAGAGA 315

Qy 652 ATCACTTAAACCCAGGAGGTGGAGGTTGCAGTGAGCTGAGATCTGTCGCACT-CACTCAA 710

Db 314 ATTACTTGAACCTGGGAGGTGGAGGTTGCAGTGAGCCCAAGATTGCACCATCTACACTCCAG 255

Qy 711 CCTGGGAGACAGAGTGACACTTTTGTCTCAAAAAAGAAAAAAGAAAAAAGAAAAAGAA 770

Db 254 CCTGGGAGACAAAGTGAGAC-TCCGTCTCAAAAAAAGAAAAAAGAAAAAAGAAAAAGAG 196

Qy 771 ACAGGTGAAGTTAACTTTAA 790

Db 195 AACTTTCAGTTACCTAGAA 176

RESULT 8  
CD518470

LOCUS CD518470

DEFINITION AGENCOURT 14368329 NIH MGC 181 Homo sapiens cDNA clone IMAGE:30396254 5', mRNA sequence.

ACCESSION CD518470

VERSION CD518470.1 GI:31450188

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1019)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
CDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM477 row: k column: 15  
High quality sequence stop: 558.

FEATURES  
source  
1. .1019  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30396254"  
/tissue\_type="White Matter"  
/dev\_stage="Unknown"  
/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances) "  
/clone\_lib="NIH MGC 181"  
/note="vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 21.6%; Score 216; DB 6; Length 1019;  
Best Local Similarity 73.2%; Pred. No. 1.9e-23;  
Matches 287; Conservative 0; Mismatches 104; Indels 1; Gaps 1;

Qy 433 AGCCACATTAACAGGTAAAGGCTGGCGCAGTGGCTCACACCTGTAAATCCAGCA 492

Db 221 ACCCAGATTAAGAGATTGAAGAGCCCGCAGTGGCTCATGCTGTAAATCCAGCA 280



/lab host="DH10B"  
/clone lib="686" (synonym: hlcc3)"  
/note="Vector: triplex2; Site\_1: SfiI; Site\_2: SfiI; B;  
CDNA-collection"

ORIGIN

Query Match 21.5%; Score 215.4; DB 5; Length 573;  
Best Local Similarity 82.0%; Pred. No. 2.9e-23;  
Matches 260; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 464 CGCAGTGGCTCACACCTTAATCCAGCATTGGAGGCTGAGCGAGATCACCTT 523  
Db 1 CGCAGTGGCTCACACCTTAATCCAGCATTGGAGGCTGAGCGAGATCACCTG 60

Qy 524 TGGTCAGAGTTGAGATAGCTGGCCAAATGCGGAACTCTGTCTCTAAAAAAT 583  
Db 61 AGGTGAGAGTTCAAGATAGCTGGCCAAATGCGGAACTCTGTCTCTAAAAA 120

Qy 584 AAAAAATTAGCTGGCATGTGGCGGCGCTGTAACTCAGCTGCTCAGGCGCGAG 643  
Db 121 AAAAAATTAGCTGGCATGTGGCGGCGCTGTAACTCAGCTGCTCAGGCGCGT 180

Qy 644 ACAAGAATCACTTAACCCAGGAGTGGAGTTGCACTGAGTGGTGCCTCT- 702  
Db 181 CGCAGAGATCGCTTGAACCCAGGAGTGGAGTTGCACTGAGTGGTGCCTCT 240

Qy 703 CACTCCACCTGGAGACAGATGACATTTTGTCTCAAAAGAAAAAACAAGTA 762  
Db 241 CACTCCACCTGGGCGACAGATCGAGTCTGTCTCAAAAGAAAAAAGAA 300

Qy 763 AAAAGAAACAGGTGAA 779  
Db 301 GAAACAGGAGAGGA 317

RESULT 11  
CD101954/c

LOCUS CD101954 823 bp mRNA linear EST 15-MAY-2003  
DEFINITION AGENCOURT 13980468 NIH MGC 187 Homo sapiens cDNA clone  
IMAGE:30373404 5', mRNA sequence.

ACCESSION CD101954  
VERSION CD101954.1 GI:30755128  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 823)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution by: Agencourt Bioscience Corporation  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCM167 row: c column: 13  
High quality sequence stop: 327.  
Location/Qualifiers  
1. 823  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30373404"  
/lab host="DH10B (T1 phage-resistant)"  
/clone lib="NIH MGC 187"  
/note="Organ: Blood vessels - aorta, basilar and artery;  
Vector: pDNR-LiB; Site\_1: SfiI (ggccattatggcc); Site\_2:  
SfiI (ggccgctcgcc); 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:

5'-CACGCCATTATGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGGCGGCGGCGGCACATG-DT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4 kb  
(range 0.5-4.0 kb). 14/15 colonies contained inserts by  
PCR. This library was enriched for full-length clones and  
was constructed by Clontech Laboratories (Palo Alto, CA).  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 21.5%; Score 215.4; DB 6; Length 823;  
Best Local Similarity 80.9%; Pred. No. 2.6e-23;  
Matches 263; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

Qy 444 AACAGGTAAAGAGCTGGGCGCAGTGGCTCACACCTGTATATCCAGCACATTTGGAGGC 503  
Db 326 AAGATGAATGAGCGCCAGGTAAAGTGGCTCATACCTGTATATCCAGCACATTTGGAGGC 267

Qy 504 TGAGCGAGGAGATCAGCTTTGGTCAGGAGTTTGGAGTACCTGGCCAAATGCGGAAA 563  
Db 266 TGAGCGAGGAGATCAGCTTTGGTCAGGAGTTTGGAGTACCTGGCCAAATGCGGAAA 207

Qy 564 CTCTGTCTCTAAAAAATAACAAAAATAGCTTGGCATGTGGCGGCGCTGTAACTCT 623  
Db 206 CCCTGTCTCT-ATAAATAATACAAAAATAGCTTGGTCAGTGGCGGCTGTAACTCT 148

Qy 624 CAGCTGCTCAGGAGGCGGAGACACAGAAATCACTTAAACCCAGGAGTGGAGTTGCACT 683  
Db 147 CTGCTACTTGGAGGCTGAGACAGAGAAATCACTTGAACCCAGGAGGTAGAGTTGCACT 88

Qy 684 GAGCTGAGATCGTCCACTCACTCACTTGGGAGACAGATGACATTTTGTCTCAAAA 743  
Db 87 GAGCTGAGATCATGCCACTGCCTCAGCTGGGCAACAGAGCAAACTCTGTCTCAAAA 28

Qy 744 AGAAAAAATAACAAAGTAAAAAAG 768  
Db 27 AAAAAAATAAATAAATAAATAAATG 3

RESULT 12  
AV759632

LOCUS AV759632 397 bp mRNA linear EST 19-OCT-2000  
DEFINITION MDS Homo sapiens cDNA clone MDSDBF02 5', mRNA sequence.  
ACCESSION AV759632  
VERSION AV759632.1 GI:10917480  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 397)  
AUTHORS Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,  
Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,  
Zeng, B., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,  
Lu, G., Yang, Y., Gao, G., Zhang, Q., Chen, S., Han, Z. and Chen, Z.  
TITLE Homo sapiens cDNA MDS clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers  
1. 397  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MDSDBF02"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"





/lab host="DH10B (phage-resistant)"  
 /clone lib="Human insulinoma"  
 /note="Organ: pancreas; Vector: pBluescript SK-; Site 1:  
 XhoI; Site 2: EcoRI; Constructed with lambda ZAPII system  
 (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to  
 pBluescript SK- by Dr. H. Inoue following the Washington  
 University protocol  
 (http://genome.wustl.edu/est/lambda\_protocol.shtml).  
 Please contact Hiroshi Inoue, MD/PhD for further  
 information on this library (Metabolism Division, Permutt  
 Laboratory, Washington University School of Medicine, Box  
 8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this  
 is a Washington University Pancreas EST project library."

## ORIGIN

Query Match 21.5%; Score 215; DB 7; Length 633;  
 Best Local Similarity 83.4%; Pred. No. 3.2e-23;  
 Matches 256; Conservative 0; Mismatches 50; Indels 1; Gaps 1;  
 QY 454 AAAGCGTGGCGCAGTGGCTCACACCTGTAATCCAGCACATTTGGGAGGCTGAGCGCAGGC 513  
 Db 308 ACAGCGCGGCGCAGTGGCTCACCGCTGTAATCCAGCATTTTGGGAGGCGGAGCAGGT 249  
 QY 514 AGATCACCTTTGGTCAGGAGTTTGAGACTGACCTGGCCCAACATGCGCAAACTCTGTCTCT 573  
 Db 248 GGATCACCTGAGGTGAGGAGTTTGAGACCACTGGCCCAACATGCTGTAACCCGCTCT 189  
 QY 574 AAAAAAATACAAAATAGCCTGGCATGATGGCGGCGCTGTAACTCAGCTGCTCA 633  
 Db 188 ACTAAAAAATACAAAATAGCCTGGCATGATGGCGGCGCTGTAACTCAGCTGCTCA 129  
 QY 634 GGAGCGGAGACACAGAAATCACTAAACCCAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 693  
 Db 128 GGAGCGTGGAGCAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 69  
 QY 694 CGTGCCACT-CACCTCCAACTGGGAGACAGAGTGCACATTTTGTCTCAAAAAGAAAAA 752  
 Db 68 CACACCTTGCATCCAGCCTGGTGGACAGAGTGCAGTCTGTCTCAAGAGGAAAAA 9  
 QY 753 AAAACAA 759  
 Db 8 AAAAAA 2

## RESULT 15

AW516097/c  
 LOCUS  
 DEFINITION xt60f05.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2790849 3,  
 similar to contains 'Alu repetitive element', mRNA sequence.  
 ACCSSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 491)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapsb-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 image.llnl.gov/image/html/iresources.shtml  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 424.

## FEATURES

source  
 1..491  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2790849"  
 /tissue\_types="moderately-differentiated endometrial  
 adenocarcinoma, 3 pooled tumors"  
 /lab host="DH10B"  
 /clone\_lib="NCI-CGAP\_Ut2"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.85 kb. Life Technologies catalog #:  
 11539-012"

## ORIGIN

Query Match 21.4%; Score 214; DB 2; Length 491;  
 Best Local Similarity 79.4%; Pred. No. 5e-23;  
 Matches 278; Conservative 0; Mismatches 70; Indels 2; Gaps 2;  
 QY 428 TTAGTAGCCACATTAACACAGGTAAAAAGCTGGGCGCAGTGGCTCACACCTGTAATCC 487  
 Db 481 TTAGTGTATTATCTTTGAAAAACATGTACTTAGGCTGGGCGCGTGGCTCACGCTGTAATCC 422  
 QY 488 CAGCACCTTTGGAGGCTGAGGAGGAGTCCACCTTTTGGTCAGGAGTTTGAGACTAGCCT 547  
 Db 421 CAGCACCTTTGGAGGCTGAGGAGGAGTCCACCTGAGTGGAGTTTGAGACAGCCT 362  
 QY 548 GGCCAAACATGGCGAACTCTGTCTCTAAAAAATAACAAAAATAGCCTGGCATGGTGG 607  
 Db 361 GACCAAAAAGGTCAAACTCCGTCTCT-CTAAAAAATACAAAAATGGCCAGCATGGTAG 303  
 QY 608 CGGCGCGCTGTAATCTCAGCTGCTCAGGAGGCCAGACACAAATCACTTAAACCCAGG 667  
 Db 302 CAGACGCTGTAGTTCAGCTATTTCAGGAGGATGAGACAGGAGAAATTGCTTGAACCCAGG 243  
 QY 668 AGGTGGAGGTGGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 726  
 Db 242 AGGTGGAGGTGGCAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 183  
 QY 727 ACACCTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGGT 776  
 Db 182 AGACTCTTTGTCTCAAAAAGAAAAAAGAAAAAGAAAAAACAATGT 133

Search completed: July 4, 2005, 15:07:07  
 Job time : 2183.56 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 18:09:16 ; Search time 2885.07 Seconds

(without alignments)  
16811.954 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_8000\_9000

Perfect score: 1001

Sequence: 1 aacagaagaagcacatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hhg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	6	CQ788219 Sequence
2	1001	100.0	11570	6	CQ874885 Sequence
3	1001	100.0	11570	6	CQ874961 Sequence
4	1001	100.0	11570	9	AF135028 Homo sapi
5	1001	100.0	107487	9	AC011483 Homo sapi
6	1001	100.0	230000	9	AF243527 Homo sapi
7	983.6	98.3	217346	2	AC027802 Homo sapi
8	978.6	97.8	200792	2	AC130782 Homo sapi
9	289.4	28.9	165436	9	HSAL60H22 Human DNA
10	289.4	28.9	230141	2	AL359177 Homo sapi
11	286	28.6	129866	9	AL138703 Homo sapi
12	285.4	28.5	230141	2	AL359177 Homo sapi
13	285	28.5	194197	9	AC064871 Homo sapi
14	284.2	28.4	98868	9	AL139231 Human DNA
15	283.4	28.3	143947	2	AC012462 Homo sapi
16	282.8	28.3	168686	2	AC084008 Homo sapi
17	282.6	28.2	172747	9	AL359835 Human DNA
18	282.4	28.2	26590	9	AC090672 Homo sapi
19	282.4	28.2	71506	9	AC092452 Homo sapi

C 20	282.4	28.2	162712	2	AC026096 Homo sapi
C 21	282.4	28.2	206533	9	CNS01RH1 Human chr
C 22	282.2	28.2	59257	9	AC106019 Homo sapi
C 23	282.2	28.2	65545	2	AC139371 Homo sapi
C 24	282.2	28.2	164282	2	AC015734 Homo sapi
C 25	282.2	28.2	196037	9	AC129492 Homo sapi
C 26	281.8	28.2	174654	9	AC073167 Homo sapi
C 27	281.6	28.1	166459	9	AC016092 Homo sapi
C 28	281.2	28.1	129686	9	HS01824H1 Human DNA
C 29	281.2	28.1	149445	9	AC013264 Homo sapi
C 30	281	28.1	173844	9	AC078842 Homo sapi
C 31	280.8	28.1	224573	9	AC008758 Homo sapi
C 32	280.6	28.0	152457	9	AC122693 Homo sapi
C 33	280.6	28.0	162401	2	AC024024 Homo sapi
C 34	280.6	28.0	172272	9	AC027804 Homo sapi
C 35	280.4	28.0	144683	9	AC096554 Homo sapi
C 36	280.4	28.0	146443	9	AC007488 Homo sapi
C 37	280.4	28.0	172092	2	AC013341 Homo sapi
C 38	280.4	28.0	184439	2	AP003085 Homo sapi
C 39	280.2	28.0	155217	9	AC105343 Homo sapi
C 40	280.2	28.0	161010	2	AC146128 Pan trogl
C 41	280.2	28.0	168203	2	AP001902 Homo sapi
C 42	280	28.0	186113	9	AC129097 Papio anu
C 43	280	28.0	189315	9	AC098848 Homo sapi
C 44	279.8	28.0	80658	9	AL137021 Human DNA
C 45	279.8	28.0	159934	9	AL160273 Human DNA

ALIGNMENTS

RESULT 1					
CQ788219	CQ788219	Sequence 2 from Patent WO2004021008.	11570 bp	DNA	linear PAT 24-MAR-2004
LOCUS	CQ788219				
DEFINITION	CQ788219				
ACCESSION	CQ788219				
VERSION	CQ788219.1	GI:45723068			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1				
AUTHORS	Diamandis, E.P.				
TITLE	Methods for detecting breast and ovarian cancer				
JOURNAL	Patent: WO 2004021008-A 2 11-MAR-2004;				
FEATURES	MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)				
LOCATION/Qualifiers					
source	1..11570				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN					
Query Match	100.0%	Score 1001;	DB 6;	Length 11570;	
Best Local Similarity	100.0%	Pred. No. 2.6e-246;			
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	AACAGAGAAAGCACATCTCGCGCGGGCGTGGCTCACACCTGTATATCCCAACACTTT	60		
DB	8000	AACAGAGAAAGCACATCTCGCGCGGGCGTGGCTCACACCTGTATATCCCAACACTTT	8059		
QY	61	GGGAGGCGAGCGGGGTGGATCCTAGGTACAGGAGATGAGACCATCTCTGGCTGACACGG	120		
DB	8060	GGGAGGCGAGCGGGGTGGATCCTAGGTACAGGAGATGAGACCATCTCTGGCTGACACGG	8119		
QY	121	TGAACCTCTCTCTACTATAAACACAAAAATTAGCCGGCGGTGGTGGCGGCGCTGT	180		
DB	8120	TGAACCTCTCTCTACTATAAACACAAAAATTAGCCGGCGGTGGTGGCGGCGCTGT	8179		
QY	181	AGTCCCACTACTCTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGAGCGGAACCTT	240		
DB	8180	AGTCCCACTACTCTCGGAGGCTGAGGAGGAGATGGCTGAACCCGGAGCGGAACCTT	8239		

Qy	241	GCAGTGAGCCGAGGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCCGTCT	300
Ds	8240	GCAGTGAGCCGAGGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCCGTCT	8299
Qy	301	CAAAAAAAAAAAGAAAAAGAAATCACTCTCAATCAAGTGGTGGCAATTTAAAC	360
Ds	8300	CAAAAAAAAAAAGAAAAAGAAATCACTCTCAATCAAGTGGTGGCAATTTAAAC	8359
Qy	361	TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTGTGTTTTCCAGACCTCAAGGTTGTTTT	420
Ds	8360	TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTGTGTTTTCCAGACCTCAAGGTTGTTTT	8419
Qy	421	TTGTTTGTGTTTTTATACCGGTGTGCTGTGGGTGGCCACTAAAAAGCTTACAAGCAAG	480
Ds	8420	TTGTTTGTGTTTTTATACCGGTGTGCTGTGGGTGGCCACTTAAAGCTTACAAGCAAG	8479
Qy	481	AAATAATAACAACATAACAATACTAATAACAATAAGTATATAAAAAATAATAGCATCTGGCTA	540
Ds	8480	AAATAATAACAACATAACAATACTAATAACAATAAGTATATAAAAAATAATAGCATCTGGCTA	8539
Qy	541	ATTGCTGACACTGTTTTAAGTGTGTTGCATGCCCTCAGCTCATTAACATTTACCTGTT	600
Ds	8540	ATTGCTGACACTGTTTTAAGTGTGTTGCATGCCCTCAGCTCATTAACATTTACCTGTT	8599
Qy	601	ATTATTGGCCCTTATTTTACAACCAAGGAGCCAAAGGCTCAGAGCAGTTTAACTTAAACGCTC	660
Ds	8600	ATTATTGGCCCTTATTTTACAACCAAGGAGCCAAAGGCTCAGAGCAGTTTAACTTAAACGCTC	8659
Qy	661	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTTAAACCAA	720
Ds	8660	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTTAAACCAA	8719
Qy	721	GAAAGTTGAAATTTAGAGGTACAGGCAAGCTTAAGCTTGTGTTTTTGAACAGTGTCTGCT	780
Ds	8720	GAAAGTTGAAATTTAGAGGTACAGGCAAGCTTAAGCTTGTGTTTTTGAACAGTGTCTGCT	8779
Qy	781	ACTGGAAAAAGGCAAGTCTTGGCTTCTCTAATAATGATACCAGGACTCTGTAATTCAT	840
Ds	8780	ACTGGAAAAAGGCAAGTCTTGGCTTCTCTAATAATGATACCAGGACTCTGTAATTCAT	8839
Qy	841	ATTTTGCATGATGTAAGTAAAGAAATGAAGCGGGTGCAATGGCACATGCCAGTAATCCC	900
Ds	8840	ATTTTGCATGATGTAAGTAAAGAAATGAAGCGGGTGCAATGGCACATGCCAGTAATCCC	8899
Qy	901	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCTG	960
Ds	8900	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCTG	8959
Qy	961	GGCAACTAAAAATTTAAAAAATAAAAAATACTAAATTTGTTTTT	1001
Ds	8960	GGCAACTAAAAATTTAAAAAATAAAAAATACTAAATTTGTTTTT	9000

RESULT 2  
CO874885  
LOCUS CO874885 11570 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 3 from Patent WO2004077060.  
ACCESSION CO874885  
VERSION CO874885.1 GI:52748035  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
AUTHORS Diamandis E.P. and Petraki C.  
TITLE Assay for detection of renal cell carcinoma  
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;  
Mount Sinai Hospital (CA)  
FEATURES  
1..11570  
Location/Qualifiers  
source /organism="Homo sapiens"

		/mol_type="unassigned DNA"		/db_xref="taxon:9606"	
ORIGIN		Query Match 100.0%; Score 1001; DB 6; Length 11570; Best Local Similarity 100.0%; Pred No. 2.6e-246; Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	AACAGAGAAAGCACATCTCGCGCGGGCGGTGGCTCACACCTGTAAATCCCAACACTTTT		60	
Db	8000	AACAGAGAAAGCACATCTCGCGCGGGCGGTGGCTCACACCTGTAAATCCCAACACTTTT		8059	
Qy	61	GGGAGGCCGAGGGGTGGATCACTAGGTCAGAGATGAGAGACCATCTGGCTGACACGG		120	
Db	8060	GGGAGGCCGAGGGGTGGATCACTAGGTCAGAGATGAGAGACCATCTGGCTGACACGG		8119	
Qy	121	TGAAACCTCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGCGCCTGT		180	
Db	8120	TGAAACCTCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGCGCCTGT		8179	
Qy	181	AGTCCCACTACTCTCGGGAGGCTGAGGCGAGGAAATGGCTCTGAACCCGGGAGGCGGAACCTT		240	
Db	8180	AGTCCCACTACTCTCGGGAGGCTGAGGCGAGGAAATGGCTCTGAACCCGGGAGGCGGAACCTT		8239	
Qy	241	GCAGTGAGCGAGGTTGCAACCACTGCATCTCAGCTGGGCAACACAGTGAGACTCCGCTCT		300	
Db	8240	GCAGTGAGCGAGGTTGCAACCACTGCATCTCAGCTGGGCAACACAGTGAGACTCCGCTCT		8299	
Qy	301	CAAAAAAAAAAAGAAAAAGAAAGAAATCACATCTCATTTCAAGTGTGGCATTTAAAC		360	
Db	8300	CAAAAAAAAAAAGAAAAAGAAAGAAATCACATCTCATTTCAAGTGTGGCATTTAAAC		8359	
Qy	361	TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTGTGTTTTTCCAGACTCAAGTGTTGTTTT		420	
Db	8360	TATTTAGCCTTTCTGTAGGCAAGTTAGTATCTTGTGTTTTTCCAGACTCAAGTGTTGTTTT		8419	
Qy	421	TTGTTTGTGTTTTTCAATACCGGTGTGTGGTCTGGGTGGCCACTATAAAGCTACAAGCAAG		480	
Db	8420	TTGTTTGTGTTTTTCAATACCGGTGTGTGGTCTGGGTGGCCACTATAAAGCTACAAGCAAG		8479	
Qy	481	AAATAATAACAACATAACAATACTAATAACCAATAGTATATAAAAAATAATAGCATCTGGCTA		540	
Db	8480	AAATAATAACAACATAACAATACTAATAACCAATAGTATATAAAAAATAATAGCATCTGGCTA		8539	
Qy	541	ATTGCTGGACACTGTTTTTAAAGTGTGTTTGCATGCTCAGCTCATTTAACTCATTTACCTGTT		600	
Db	8540	ATTGCTGGACACTGTTTTTAAAGTGTGTTTGCATGCTCAGCTCATTTAACTCATTTACCTGTT		8599	
Qy	601	ATTATTGGCCCTATTTTACAACCAAGGAGCCAAAGGCTCAGAGCAGTTAACTAACAGCCTC		660	
Db	8600	ATTATTGGCCCTATTTTACAACCAAGGAGCCAAAGGCTCAGAGCAGTTAACTAACAGCCTC		8659	
Qy	661	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTTAAACCAAA		720	
Db	8660	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTTAAACCAAA		8719	
Qy	721	GAAAGTTGAAATTTAGAGGTACAGGCAAGCTTAAGCTTGTGTTTTTGAACAGTGTCTGCT		780	
Db	8720	GAAAGTTGAAATTTAGAGGTACAGGCAAGCTTAAGCTTGTGTTTTTGAACAGTGTCTGCT		8779	
Qy	781	ACTGGAAAAAGGCAAGTCTTGGCTTTCTCTAATAATTGATACCAGGACTCTGTAATTCAT		840	
Db	8780	ACTGGAAAAAGGCAAGTCTTGGCTTTCTCTAATAATTGATACCAGGACTCTGTAATTCAT		8839	
Qy	841	ATTTTGCATGATGTAAGTAAAGAAATGAAGCGGGGTGCAATGGCACATGCCAGTAATCCC		900	
Db	8840	ATTTTGCATGATGTAAGTAAAGAAATGAAGCGGGGTGCAATGGCACATGCCAGTAATCCC		8899	
Qy	901	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCCTG		960	
Db	8900	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCCTG		8959	
Qy	961	GGCAACTAAAAATTTAAAAAATAAAAAATACTAAATTTGTTTTT		1001	



/db\_xref="GI:4589283"  
translation="MATARPPEWMMVLICALITALLLGVTEHVLANNNDVSCDHPNSTVPS  
GNSDLGAGEDARSDDSSRIINGSDMDHTQWQAALLLRNQLYCGAVLVHPQW  
LLTAHCKKVRVRLGHYSLSPVESQDMFQGVKSIPIHPGVSHPCNSDMLIKLN  
RRIPTKDVRLNVSSHCPKSLAGTCLVSGWGTAKSPVHPKVLQCLNLSVLSQKCE  
DAYFRQDDTFNFCAGDKAGRSDSCGDSGGPVVCGSLQGLVSGDYPFARNRPGVIT  
NLCKFTKWIQETIQANS"

ORIGIN

Query Match 100.0%; Score 1001; DB 9; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 2.6e-246; Indels 0; Gaps 0;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 60  
Db 8000 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 8059  
Qy 61 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120  
Db 8060 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 8119  
Qy 121 TGAACCCCTGTCTTACTAAAAACACAAAAAATTAGCCGGCGTGGTGGAGCGCCTGT 180  
Db 8120 TGAACCCCTGTCTTACTAAAAACACAAAAAATTAGCCGGCGTGGTGGAGCGCCTGT 8179  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240  
Db 8180 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 8239  
Qy 241 GCAGTGAGCGAGGTGCGACCTGCACTCGAGCTGGGCAACACAGTGAACCTCGTCT 300  
Db 8240 GCAGTGAGCGAGGTGCGACCTGCACTCGAGCTGGGCAACACAGTGAACCTCGTCT 8299  
Qy 301 CAAAAAAGAAAAAGAAAGAAATCACTCTATTCATGCTGGTGGGCAATTAAC 360  
Db 8300 CAAAAAAGAAAAAGAAAGAAATCACTCTATTCATGCTGGTGGGCAATTAAC 8359  
Qy 361 TATTAGCTTCTGTAGGCAAGGTAGTATCTTGTGTTTCCAGACCTCAAGGTGTTT 420  
Db 8360 TATTAGCTTCTGTAGGCAAGGTAGTATCTTGTGTTTCCAGACCTCAAGGTGTTT 8419  
Qy 421 TTGTTTGTGTTTTCATACCGGTGTGTGTGGTGGTGGCCACTAAAGCTACAGCAAG 480  
Db 8420 TTGTTTGTGTTTTCATACCGGTGTGTGTGGTGGTGGCCACTAAAGCTACAGCAAG 8479  
Qy 481 AAATAATACACTACACAACTACTAATACCAATAGTATTAATAATATAGCATCTGCTA 540  
Db 8480 AAATAATACACTACACAACTACTAATACCAATAGTATTAATAATATAGCATCTGCTA 8539  
Qy 541 ATTGCTGGACACTGTTTAAAGTGTGTTGCATGCTCAGCTCATTAACCTCATTTACCTGT 600  
Db 8540 ATTGCTGGACACTGTTTAAAGTGTGTTGCATGCTCAGCTCATTAACCTCATTTACCTGT 8599  
Qy 601 ATTATGCGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCTTAACCTAAAGCTC 660  
Db 8600 ATTATGCGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCTTAACCTAAAGCTC 8659  
Qy 661 TCAGAGAACTCTGCAGAGATATTAATTTAAATAATAATAGAGAAATTAACCAAA 720  
Db 8660 TCAGAGAACTCTGCAGAGATATTAATTTAAATAATAATAGAGAAATTAACCAAA 8719  
Qy 721 GAAAGTTCAAAATTTAGAGTACAGGCTTAAGCTTGTGCTTTGAAACAGTGTCTGCT 780  
Db 8720 GAAAGTTCAAAATTTAGAGTACAGGCTTAAGCTTGTGCTTTGAAACAGTGTCTGCT 8779  
Qy 781 ACTGGAAAAAGGCAAGCTTTGGCTTTCTTAATATGATACAGGACTCTGTAATTCAT 840  
Db 8780 ACTGGAAAAAGGCAAGCTTTGGCTTTCTTAATATGATACAGGACTCTGTAATTCAT 8839  
Qy 841 ATTTTGCATGCTGTAAGTAAAGTAAAGCGGGTGAATGGCACATGCCAGTAATCCC 900  
Db 8840 ATTTTGCATGCTGTAAGTAAAGTAAAGCGGGTGAATGGCACATGCCAGTAATCCC 8899

Qy 901 AGCACTCTGGAGACTGAAGTGGAGAGATCACTTGAGCTCAGGAGTTCAAGACGAGCTG 960  
Db 8900 AGCACTCTGGAGACTGAAGTGGAGAGATCACTTGAGCTCAGGAGTTCAAGACGAGCTG 8959  
Qy 961 GGCACCTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1001  
Db 8960 GGCACCTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9000  
RESULT 5  
AC011483/c  
LOCUS AC011483 107487 bp DNA linear PRI 01-JUL-2002  
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.  
AC011483  
AC011483.7 GI:21637461  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 107487)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 107487)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 107487)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.6.  
NOTE: Shatter libraries failed to resolve dinucleotide repeat.  
Unsure number of repeat copies 64998-65494. Forced join 65015.  
Location/Qualifiers  
1. 107487  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTB-147C22"  
64998..65494  
/note="NOTE: Shatter libraries failed to resolve  
dinucleotide repeat. Unsure number of repeat copies  
64998-65494. Forced join 65015."

misc\_feature  
ORIGIN  
Query Match 100.0%; Score 1001; DB 9; Length 107487;  
Best Local Similarity 100.0%; Pred. No. 3.2e-246; Indels 0; Gaps 0;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 60  
Db 52220 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATCCCAACACTTT 52161  
Qy 61 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 120  
Db 52160 GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTGGCTGACACGG 52101  
Qy 121 TGAACCCCTGTCTTACTAAAAACACAAAAAATTAGCCGGCGTGGTGGAGCGCCTGT 180  
Db 52100 TGAACCCCTGTCTTACTAAAAACACAAAAAATTAGCCGGCGTGGTGGAGCGCCTGT 52041  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTT 240

```
Db 52040 AGTCCAGCTACTCGGAGGCTGAGGCAGGAGATGGCTGAAACCGGAGGCGGAACCTT 51981
Qy 241 GCAGTGAGCCGAGGTTGCACCACTGCACCTCAGCCTGGGCAACACAGTGAGACTCCGCTCT 300
Db 51980 GCAGTGAGCCGAGGTTGCACCACTGCACCTCAGCCTGGGCAACACAGTGAGACTCCGCTCT 51921
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 360
Db 51920 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 51861
Qy 361 TATTAGGCTTTCTGTAGGCAAGGTTAGTATCTTGTGTTTTTCCAGACCTCAAGGTTGTTTTT 420
Db 51860 TATTAGGCTTTCTGTAGGCAAGGTTAGTATCTTGTGTTTTTCCAGACCTCAAGGTTGTTTTT 51801
Qy 421 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGTGGGTGGGCCCACTAAAGCTTACAAGCAAG 480
Db 51800 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGTGGGTGGGCCCACTAAAGCTTACAAGCAAG 51741
Qy 481 AAAAAAATAAACAATAACAATACTAATAACCAATAGTATTAATAATAATAAGTATAGCTGCTA 540
Db 51740 AAAAAAATAAACAATAACAATACTAATAACCAATAGTATTAATAATAATAAGTATAGCTGCTA 51681
Qy 541 ATTGCTGACACTGTTTTTAAGTGGTTTGCATGCTCAGCTCATTAACTCATTTTACCTGTT 600
Db 51680 ATTGCTGACACTGTTTTTAAGTGGTTTGCATGCTCAGCTCATTAACTCATTTTACCTGTT 51621
Qy 601 ATTATGCGCCCTATTTTACAAACAGGAGGCAAGGCTCAGAGCTTAACCTAAACAGCCTC 660
Db 51620 ATTATGCGCCCTATTTTACAAACAGGAGGCAAGGCTCAGAGCTTAACCTAAACAGCCTC 51561
Qy 661 TCAAAAGAACTCTGCAGAGATATTAATTTAAAAAATAAGTAAAGAGAAATTAACCAACAA 720
Db 51560 TCAAAAGAACTCTGCAGAGATATTAATTTAAAAAATAAGTAAAGAGAAATTAACCAACAA 51501
Qy 721 GAAAGTTGAAATTTAGAGGTACAGGACCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 780
Db 51500 GAAAGTTGAAATTTAGAGGTACAGGACCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 51441
Qy 781 ACTGGGAAAAAGGCAAGCTTTGGCTTCTTAATAATGATACCAAGGACTCTGTAAATCAT 840
Db 51440 ACTGGGAAAAAGGCAAGCTTTGGCTTCTTAATAATGATACCAAGGACTCTGTAAATCAT 51381
Qy 841 ATTTGCTATGATGATAGTAAGATGAAGCGGGTGAATGCGACATGCCAGTAATCCC 900
Db 51380 ATTTGCTATGATGATAGTAAGATGAAGCGGGTGAATGCGACATGCCAGTAATCCC 51321
Qy 901 AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCTG 960
Db 51320 AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGAGCTCAGGAGTTCAAGACCAAGCTG 51261
Qy 961 GGCACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1001
Db 51260 GGCACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 51220
```

## RESULT 6

```
AF243527/c AF243527 230000 bp DNA linear PRI 21-NOV-2000
LOCUS AF243527 Homo sapiens serine protease gene cluster, complete sequence.
DEFINITION AF243527
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 230000)
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL Gene 257 (1), 119-130 (2000)
```

```
MEDLINE 20510030
PUBMED 11054574
REFERENCE 2 (bases 1 to 230000)
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
Location/Qualifiers
1. 230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
complement(<6963..>11517)
/gene="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..>11517))
/gene="KLK1"
product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
note="serine protease"
codon_start=1
product="renal kallikrein"
protein_id="AAG33353.1"
db_xref="GI:11244758"
translation="MWFLVLCALSLGCTGAAPIQSRIIVGVGECEHOSQPQWQAALYH
FSYFQCGGIIIVHQWVLTAAHCISDNYQLMGRHLPDDTAQFVHVSFPIHGFEN
MSLEHTRQADEYSHDLMLRTEPADTITDAVKVVELPTEPEVSTFCIAGSWG
IIEPENSFPDQLQCDLIKILPNEKSKAVKQVDFMLCVHLEGGKDTGVDGSGGL
MCDGLQGVTSWGVPGCTPKNSKAVRVLVSYVKWIEDTTAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
product="ACO protease"
complement(join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..19246))
note="serine protease"
codon_start=1
product="ACO protease"
protein_id="AAG33354.1"
db_xref="GI:11244759"
translation="MWLLTLFLSLASTAAQDGKLLLEGDECAPHSQPQWQVALYERGR
FNCASLISPHWVLSAHCQSRFRVRLGSHNLKRDGPQLRTSRVIPHRYEARS
HRNDIMLRVQPARLNPOVRPAVLPTFCPHGECACVVGWGLVPLSPVSLPDTLHC
ANTISLSDTSCDKSYPCGLRNTMVCAGAEGRGAESCDSGGLVCGGILQGIWSGD
VPCDNTKPGVYKVCVHYLEWIRETMKRN"
<42595..>47769
/gene="KLK3"
join(<42595..42640,43880..44039,45669..45955,46099..46235,
47614..>47769)
/gene="KLK3"
product="prostate specific antigen"
join(42595..42640,43880..44039,45669..45955,46099..46235,
47614..47769)
/gene="KLK3"
note="serine protease"
codon_start=1
product="prostate specific antigen"
protein_id="AAG33355.1"
db_xref="GI:11244760"
translation="MWVTVVFLTSLVTWIGAAPLISRIIVGVGECEKHOSQPQWVLVAS
RGAVCGGLVHPQWVLTAAHCIRNKSIVLLGRHSLPHEDTGVGVFVSHSFPHPPLVD
MSLLKRPRLRPDDSHDLMLRLSEPAELTDAVKMDLPTOEPALETGTCYASGWSGI
EPBEFLTPKKLQCVDLINSDVCAQVHPQKVKFMLCAGRWTKGKSTCGSDSGGLV
CNGVLQGITSGSEPCALPERPSLYTKVHYRWIKDTIVANP"
<61139..>66229
/gene="KLK2"
join(<61139..61184,62391..62550,64142..64428,64542..64678,
66074..>66229)
/gene="KLK2"
```



```
CDS      /product="glandular kallikrein"
join(61139..61184,62391..62550,64142..64428,64542..64678,
66074..66229)
/gene="KLK2"
/note="serine protease"
/codon_start=1
/product="glandular kallikrein"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translacion="MWDLVLSIALSVGCTGAVPLIOSRIVIGWMECEKHSOPVOAVVS
HGWAHCGVLHPQVLTAAHCLKNSQVWLGRHNLPEPDTGQRVPVSHSPHPLIN
MSLKHQSLRPDESDDLMLLRUSEPAKITDVVKVLGLPTQEPALGTTCYASGWSGI
EPEFLRPSLQCVSLHLLSNDMCARAYSEKVTFFMLCAGLTGGDKTCGDSGGPLV
CNGVLQITSGPCCALPEKPAVYTKVHYRWIKDTIAANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/product="serine protease"
complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..98405))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translacion="MATAGNPPWGLFYLIVGAGSLVSGSCSOIINGDCSPHSQPW
QAALVMEELFCSGVLHPQVLSAAHCFQNSYITIGLHSLLEADQEPGSGQVSEASLS
VHPHYNRPILLANDMLIKLDESVESEDITRSISIAQCPTAGNSCLVSGWGLIANGR
MPTVLCQVNVSWVEECSKLYDPLYPHPSMFCAGGHDQKDCNGDSGGPLICNGYILQ
GLVSFGAPCGQVPGVYTNLCKFTWIEKTQAS"
complement(<131301..>140303)
/gene="KLK5"
/note="synonym: SCTE"
complement(join(<131301..131456,136310..136443,
136529..136785,137525..137690,140255..>140303))
/gene="KLK5"
/product="stratum corneum trypsin-like serine protease"
complement(join(131301..131456,136310..136443,
136529..136785,137525..137690,140255..140303))
/gene="KLK5"
/codon_start=1
/product="stratum corneum trypsin-like serine protease"
/protein_id="AAG33358.1"
/db_xref="GI:11244763"
/translacion="MWLICALITALLGLVTDARSDDSSRIINGSDCDHTQPWQAAL
LLRNOLYCGAVLWPLTAACHKKVFRVLGHYSLSGFVBSGQOMQGVKSIHP
PGYSHPGSNDMLIKLNRIRPTKDVPIINVSHCPISAGTKCLVSGWGTTSQVHF
PKVLQCLNISVLSKREDAYPROIDTMECAGKAGRDCQSGSPVVCNGSLQGL
VSWGDPICARNRPGVYTNLCKFTWIEKTQAS"
complement(join(<146834..146986,149628..149764,
151186..151433,155052..155208,155948..>155987))
/product="protease M"
complement(join(146834..146986,149628..149764,
151186..151433,155052..155208,155948..155987))
/note="serine protease; also called neurosin or zyme"
/codon_start=1
/product="protease M"
/protein_id="AAG33359.1"
/db_xref="GI:11244764"
/translacion="MKKLMMVLSLIAAAWAEQNKLVHGGPCDKTSHPYQAALYTSGH
LLCGVLHPUNVLTAAHCKPNLQVFLGKINLRQRESSQSSVRAVHPDIDAAS
HDQIMLLRLARPLKLSLIQPLERDCSANTTSCHILGWKGTADGFPDTIQAYI
HVRECEHAYPQGITQNMLCAGDEKYKDCSCQDSGGPLVCGDHLRLGLVSWGNIPC
GSKKPGVYTNVCRVYTNWIKTIQAK"
complement(<165420..>170283)
/gene="KLK7"
/note="synonym: SCTE"
complement(join(<165420..165575,167672..167808,
168124..168371,169651..169798,170211..>170283))
/gene="KLK7"
/product="stratum corneum chymotryptic enzyme"
complement(join(165420..165575,167672..167808,
168124..168371,169651..169798,170211..170283))
/gene="KLK7"
/note="serine protease"
Qy      1  AACAGAGAAAGACATCTCGCGCGCGGTGGCTCACACCTGTATCCCAACACTTT 60
Db      134548 AACAGAGAAAGACATCTCGCGCGCGGTGGCTCACACCTGTATCCCAACACTTT 134489
Qy      61  GGGAGGCCGAGCGGGTGGATCACTAGGTGAGGAGATGGAGACCATCTGGCTGACACGG 120
Db      134488 GGGAGGCCGAGCGGGTGGATCACTAGGTGAGGAGATGGAGACCATCTGGCTGACACGG 134429
Qy      121  TGAACCCCTGTCTTACTATAAACAACAAAAATTAGCCGGCGGTGGTGGAGGCGCCTGT 180
Db      134428 TGAACCCCTGTCTTACTATAAACAACAAAAATTAGCCGGCGGTGGTGGAGGCGCCTGT 134369
Qy      181  AGTCCCAGCTACTCGGAGGCTGAGGCGAGGAGATGCCTGAACCCGGAGGCGCGAACTT 240
Db      134368 AGTCCCAGCTACTCGGAGGCTGAGGCGAGGAGATGCCTGAACCCGGAGGCGCGAACTT 134309
Qy      241  GCAGTGAGCGGAGTTGCACCATCTGCACCTCGAGCTCGGGCAACACAGTGAGACTCCGCT 300
Db      134308 GCAGTGAGCGGAGTTGCACCATCTGCACCTCGAGCTCGGGCAACACAGTGAGACTCCGCT 134249
Qy      301  CAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGGTGGCAATTAAAC 360
Db      134248 CAAAAAAGAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGGTGGCAATTAAAC 134189
Qy      361  TATTAGCCTTCTGTAGGCAAGTTAGTATCTTGTGTTTTCCAGACCTCAAGGTGTTTTT 420
Db      134188 TATTAGCCTTCTGTAGGCAAGTTAGTATCTTGTGTTTTCCAGACCTCAAGGTGTTTTT 134129
Qy      421  TTGTTGTTTTTTCATACCCGTGTGTGTGTGGGTGGGCCACTAAAAAGCTACAAGCAAG 480
```

```
/codon_start=1
/product="stratum corneum chymotryptic enzyme"
/protein_id="AAG33360.1"
/db_xref="GI:11244765"
/translacion="MARSLLLPLQIILLLSLETAGEEAQDKIIDGAPCARGSHPNQ
VALLSGNQHCGGVLVNERVLTAAHCKMNEYTVHLGSDTLGDRRAQRIKASKSFRHP
GYSTQTHVNDLMLVNLNSQARLSMVKVLPSPRCEPPGTCTVSGMGTTTSPDVTFF
SDLMCVDFKLISPDQCTKVYKDLLENMCLCAGIPDSKKNAACNGSDGGLVCRGTLOGL
VSMGTFCGPNPDPGVYTOVCKFTKWINDTMKGR"
complement(join(<183943..184098,185635..185768,
187865..188127,188293..188452,188967..>189036))
/product="neuropsin"
complement(join(183943..184098,185635..185768,
187865..188127,188293..188452,188967..189036))
/note="serine protease; also called ovasin"
/codon_start=1
/product="neuropsin"
/protein_id="AAG33361.1"
/db_xref="GI:11244766"
/translacion="MGRPRPRAAKTWMFLLLGGAMAGHSAQEDKVLGGHECQPHSQ
PWQAALFQGOQLLCGGVLVGNVLTAAHCKPKYTVRLGDHSLQNKDGPEQETPVQV
SIHPCYNSDVEDHNDHMLLOLRDOASIGSKVKPLSLADHCTQPCQKCTVSGMGTV
TSRENPPDTLNCAEVKIFQKKCEDAYPCQITDGMVCASSSKAGATCQDGSQGLVC
DGLQGITSGSDPCGRSDRPGVYTNICRYLDWIKKLIIGSKG"
complement(join(<190980..191129,191573..191709,
194324..194589,197048..197204,197370..>197412))
/product="kallikrein-like 3"
complement(join(190980..191129,191573..191709,
194324..194589,197048..197204,197370..197412))
/note="serine protease"
/codon_start=1
/product="kallikrein-like 3"
/protein_id="AAG33362.1"
/db_xref="GI:11244767"
/translacion="MKLLGALLSLLAGHWADTRAIABECPNSQPWQAGLFLHT
RLFCGATLSDRWLLTAAHCKPKYLVRLGHEHLHWKEGPEQLFRVTFDFPFGFNKD
LSANDHNDIMLIRLPQARLSPAVQINLSQTSFPMQCLISLGSAGVSSPKALFV
TLQCANISILENKLCHWAYPGHISDSMLCAGLWEGGRGSCQDGSGLVLCVNGTLAGVV
SGNAEPCSRPRRPVAVTVSYHLDWIWMEN"
Query Match 100.0%; Score 1001; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 3,4e-246;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



*	17511	17610: gap of 100 bp
*	17611	24602: contig of 6992 bp in length
*	24603	24702: gap of 100 bp
*	24703	35434: contig of 10732 bp in length
*	35435	35534: gap of 100 bp
*	35535	124474: contig of 89940 bp in length
*	124475	124574: gap of 100 bp
*	124575	134664: contig of 10090 bp in length
*	134665	134764: gap of 100 bp
*	134765	162343: contig of 27579 bp in length
*	162344	162443: gap of 100 bp
*	162444	208917: contig of 46474 bp in length
*	208918	209017: gap of 100 bp
*	209018	217346: contig of 8329 bp in length.

## FEATURES

```

1. .217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_lib="RPC1-11 Human Male BAC"
1. .8149
/note="assembly_fragment
clone end:SP6
vector_side:left"
misc_feature
8250..9592
/note="assembly_fragment"
misc_feature
9693..10733
/note="assembly_fragment"
misc_feature
10834..13519
/note="assembly_fragment"
misc_feature
13620..17510
/note="assembly_fragment"
misc_feature
17611..24602
/note="assembly_fragment"
misc_feature
24703..35434
/note="assembly_fragment"
misc_feature
35535..124474
/note="assembly_fragment"
misc_feature
124575..134664
/note="assembly_fragment"
misc_feature
134765..162343
/note="assembly_fragment"
misc_feature
162444..208917
/note="assembly_fragment"
misc_feature
209018..217346
/note="assembly_fragment
clone end:T7
vector_side:right"

```

## ORIGIN

	Query Match	98.3%;	Score 983.6;	DB 2;	Length 217346;
	Best Local Similarity	99.5%;	Pred. No. 1e-241;		
	Matches 997;	Conservative	0;	Mismatches 4;	Indels 1; Gaps 1;
Qy	1	AACAGAGAAAGCACATCTCGCGCGGGCTGTGGCTCACACCTCTAATCCCAACACTTT	60		
Db	196413	AACAGAGAAAGCACATCTCGCGCGGGCTGTGGCTCACACCTGTATCCCAACACTTT	196354		
Qy	61	GGGAGGCGGAGCGGGTGGATCACTAGTTCAGGAGATGGAGACCATCTTGGCTGCACGG	120		
Db	196353	GGGAGGCGGAGCGGGTGGATCACTAGTTCAGGAGATGGAGACCATCTTGGCTGCACGG	196294		
Qy	121	TGAACCCCTGTCCTCTACTAATAAACACAAATAATTAGCCGGCGGTGGTGCAGGCGCCCTGT	180		
Db	196293	TGAACCCCTGTCCTCTACTAATAAACACAAATAATTAGCCGGCGGTGGTGCAGGCGCCCTGT	196234		
Qy	181	AGTCCCAGCTACTCGGGAGGCTGAGGCGAGGAATGGGCCTGAAACCCGGGAGCGGAACTT	240		
Db	196213	AGTCCCAGCTACTCGGGAGGCTGAGGCGAGGAATGGGCCTGAAACCCGGGAGCGGAACTT	196174		

Qy	241	GCAGT	AGCGAGG	TTGGCA	CACTG	CACTC	CAGCT	CGGGCA	ACACAG	TGAGACT	CGCTG	300
Db	196173	GCAGT	AGCGAGG	TTGCAC	CACTG	CACTC	CAGCT	CGGGCA	ACACAG	TGAGACT	CGCTG	196114
Qy	301	CAAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	359
Db	196113	CAAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	AAAAA	196054
Qy	360	CTATTT	AGCCCTT	CTGTAG	GCAGG	TAGTAT	CTTTG	TTTTT	TCCAG	ACCTC	CAAGGTG	419
Db	196053	CTATTT	AGCCCTT	CTGTAG	GCAGG	TAGTAT	CTTTG	TTTTT	TCCAG	ACCTC	CAAGGTG	195994
Qy	420	TTTGTT	TGTTTTT	TCATAC	CCGGT	GTGTGG	TCTGG	GTGG	CCACT	TAAAG	CTACAAG	479
Db	195993	TTTGTT	TGTTTTT	TCATAC	CCGGT	GTGTGG	TCTGG	GTGG	CCACT	TAAAG	CTACAAG	195934
Qy	480	GAATA	ATAACA	ACTTAC	CAACT	TAATAC	CAATAG	TATATA	AAAA	TATATAG	CACTCTGG	539
Db	195933	GAATA	ATAACA	ACTTAC	CAACT	TAATAC	CAATAG	TATATA	AAAA	TATATAG	CACTCTGG	195874
Qy	540	AATTG	CTGG	CACTG	TTTAA	GTGGT	TTCAT	GTCC	TAGCT	CACTTAA	CTATTTAC	599
Db	195873	AATTG	CTGG	CACTG	TTTAA	GTGGT	TTCAT	GTCC	TAGCT	CACTTAA	CTATTTAC	195814
Qy	600	TATTAT	TGGCC	TATTT	TACAA	CAAG	GAGCA	AGGCT	CAGAG	CAGT	TAACTAA	659
Db	195813	TATTAT	TGGCC	TATTT	TACAA	CAAG	GAGCA	AGGCT	CAGAG	CAGT	TAACTAA	195754
Qy	660	CTCAAA	AGAACT	CTG	CAG	AGATAT	TAAAT	TTTAA	AAAA	TATAG	AGAAA	719
Db	195753	CTCAAA	AGAACT	CTG	CAG	AGATAT	TAAAT	TTTAA	AAAA	TATAG	AGAAA	195694
Qy	720	AGAA	GTTG	AAATTTAG	AGGTAC	AGG	CAGT	TAAAG	CTTGT	TGCTTTG	AAACAG	779
Db	195693	AGAA	GTTG	AAATTTAG	AGGTAC	AGG	CAGT	TAAAG	CTTGT	TGCTTTG	AAACAG	195634
Qy	780	TACT	GGG	AAAAAG	GCAGT	CTTGG	CTTCC	TAAAT	TGATAC	CAGG	ACTCTGT	839
Db	195633	TACT	GGG	AAAAAG	GCAGT	CTTGG	CTTCC	TAAAT	TGATAC	CAGG	ACTCTGT	195574
Qy	840	TATTTT	GCATG	CATGTAA	GTAAG	AAATCA	AGCCG	GGTGC	ATGG	CACATG	CCAGT	899
Db	195573	TATTTT	GCATG	CATGTAA	GTAAG	AAATCA	AGCCG	GGTGC	ATGG	CACATG	CCAGT	195514
Qy	900	CAG	CAC	TCTGG	GAG	ACTCA	AGTGG	GGA	AGATCA	CTTTG	AGCTC	959
Db	195513	CAG	CAC	TCTGG	GAG	ACTCA	AGTGG	GGA	AGATCA	CTTTG	AGCTC	195454
Qy	960	GGG	CAACT	TAAAA	TTTTAAAA	AAAAAT	AAAAAT	AAAAAT	AAAAAT	AAAAAT	AAAAAT	1001
Db	195453	GGG	CAACT	TAAAA	TTTTAAAA	AAAAAT	AAAAAT	AAAAAT	AAAAAT	AAAAAT	AAAAAT	195412

RESULT 8	
ACI30782/c	
LOCUS	200792 bp DNA linear HTG 22-NOV-2002
DEFINITION	Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12 ordered pieces.
ACCESSION	ACI30782
VERSION	ACI30782.2 GI:35167101
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE	Pan troglodytes (chimpanzee)
ORGANISM	Pan troglodytes
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan. 1. (bases 1 to 200792) Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.C., Brinkley,C., Brooks,S., Cariaga,K., Coleman,B., Engle,J.J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J., McGuigan,C., Pearson,R., Portnov,M.E., Prasad,A.

Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,  
 Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,  
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 200792)  
 Direct Submission  
 Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 200792)  
 Direct Submission  
 Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 On Nov 22, 2002 this sequence version replaced gi:22218452.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc\_zoo@hgrl.nih.gov  
 ----- Project Information  
 Center project name: dhz  
 Center clone name: 355A20

The sequence data in this record represents an 'enhanced'  
 version of a Phase 2 submission. Specifically, the indicated  
 order and orientation of each sequence contig has been  
 established using one or more of the following: read-pair  
 data from individual subclones, overlaps with neighboring  
 clones, alignment with available reference sequence (e.g.,  
 human), and/or confirmation by PCR testing. In addition,  
 the sequence assembly is based on at least 8X average  
 coverage in Q20 bases and has been reviewed to rule out  
 gross misassemblies, the low-quality ends of sequence  
 contigs have been trimmed away, and each base is associated  
 with a Phrap-derived quality score.

----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 196656 bases at least Q40  
 Consensus quality: 197883 bases at least Q30  
 Consensus quality: 198879 bases at least Q20  
 Insert size: 215000; agarose-fp  
 Quality coverage: 9.03x in Q20 bases; agarose-fp  
 Quality coverage: 9.73x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \*  
 \* 1 28306: contig of 28306 bp in length  
 \* 28307 28406: gap of unknown length  
 \* 28407 37856: contig of 9450 bp in length  
 \* 37857 73522: contig of unknown length  
 \* 73523 73622: contig of 35566 bp in length  
 \* 73623 83567: gap of unknown length  
 \* 83568 83667: contig of 9945 bp in length  
 \* 83668 88917: gap of unknown length  
 \* 88918 125611: contig of 5150 bp in length  
 \* 125612 125711: contig of 36694 bp in length  
 \* 125712 159879: gap of unknown length  
 \* 159880 159979: contig of 34168 bp in length  
 \* 159980 174698: contig of 14719 bp in length

\* 174699 174798: gap of unknown length  
 \* 174799 186382: contig of 11584 bp in length  
 \* 186383 186483: gap of unknown length  
 \* 186483 193344: contig of 6862 bp in length  
 \* 193345 193445: gap of unknown length  
 \* 193445 193663: contig of 5919 bp in length  
 \* 193664 194633: gap of unknown length  
 \* 194633 200792: contig of 1329 bp in length.  
 FEATURES  
 source  
 1..200792  
 /organism="pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="CH251-355A20"  
 /clone\_lib="CH251"  
 1..28306  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 28407..37856  
 /note="assembly\_fragment"  
 37957..73522  
 /note="assembly\_fragment"  
 73623..83567  
 /note="assembly\_fragment"  
 83668..88917  
 /note="assembly\_fragment"  
 88918..125611  
 /note="assembly\_fragment"  
 125712..159879  
 /note="assembly\_fragment"  
 159980..174698  
 /note="assembly\_fragment"  
 174799..186382  
 /note="assembly\_fragment"  
 186483..193344  
 /note="assembly\_fragment"  
 193445..199363  
 /note="assembly\_fragment"  
 199464..200792  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right

## ORIGIN

Query Match 97.8%; Score 978.6; DB 2: Length 200792;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-240;  
 Matches 987; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 AACAGAGAAAGCACATCTCGGCCGGCGTGGCTCACACCTGTATATCCCAACACTTT 60  
 Db 71159 AACAGAGAAAGCACATCTCGGCCGGCGTGGCTCACACCTGTATATCCCAACACTTT 71100  
 QY 61 GGGAGGCCGAGCGGGTGGATCCTAGTCTAGGAGATGAGACCATCTCTGCTGACACGG 120  
 Db 71099 GGGAGGCCGAGCGGGTGGATCCTAGTCTAGGAGATGAGACCATCTCTGCTGACACGG 71040  
 QY 121 TGAACACCTGCTCTTACTTAAACACACAAAAATATAGCGCGCGTGGTGGCAGGCGCCCTGT 180  
 Db 71039 TGAACACCTGCTCTTACTTAAACACACAAAAATATAGCGCGCGTGGTGGCAGGCGCCCTGT 70980  
 QY 181 AGTCCCGAGTACTCGGAGGCTGAGGAGGAGAGATGGCTTGAACCCGGGAGCGGAACTT 240  
 Db 70979 AGTCCCGAGTACTCGGAGGCTGAGGAGGAGAGATGGCTTGAACCCGGGAGCGGAGCTT 70920  
 QY 241 GCAGTGAGCCGAGTTGCAACCACTGCCTCAGCTGGGGCAACACAGTGAAGTCCGCTCT 300  
 Db 70919 GCAGTGAGCCGAGTTGCAACCACTGCCTCAGCTGGGGCAACACAGTGAAGTCCGCTCT 70860  
 QY 301 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 360  
 Db 70859 CAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 70800

```
QY 361 TATTTAGCCTTCTGTAGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT 420
Db 70799 TATTTAGCCTTCTGTGGGCAAGGTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT 70740
QY 421 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGTGGGTGTGGCCACTAAAAGCTTACAAGCAAG 480
Db 70739 TTGTTTGTGTTTTTCATACCGGTGTGTGCTGTGGGTGTGGCCACTAAAAGCTTACAAGCAAG 70680
QY 481 AAATAATAAACACTACAAACAATCTAATPACCAATAGTATAAAAAATAATAGCATCTGGCTA 540
Db 70679 AAATAATAAACACTACAAACAATCTAATPACCAATAGTATAAAAAATAATAGCATCTGGCTA 70620
QY 541 ATTGCTGACACTGTTTTTAAGTGGTTTGCATGCTCCTCAGCTCATTAATCATTTACCTGTT 600
Db 70619 ATTGCTGACACTGTTTTTAAGTGGTTTGCATGCTCCTCAGCTCATTAATCATTTACCTGTT 70560
QY 601 ATTATTGGCCCTATTTTACAAACAGGAGGCAAGGCTCAGAGCAGTAACTTAACAGCCTC 660
Db 70559 ATTATTGGCCCTATTTTACAAACAGGAGGCAAGGCTCAGAGCAGTAACTTAACAGCCTC 70500
QY 661 TCAAAAGAAACTCTGCAGAGATATTAAATTTTAAAAAATAATGAGAGAAATTAACCCAA 720
Db 70499 TCAAAAGAAACTCTGCAGAGATATTAAATTTTAAAAAATAATGAGAGAAATTAACCCAA 70440
QY 721 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGTTTGAACAGTGTCTGCT 780
Db 70439 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGTTTGAACAGTGTCTGCT 70380
QY 781 ACTGGGAAAAGGCAAGCTTGGCTTCTCTAATTAATGATACAGGACTCTGTAATTCAT 840
Db 70379 ACTGGGAAAAGGCAAGCTTGGCTTCTCTAATTAATGATACAGGACTCTGTAATTCAT 70320
QY 841 ATTTTGCATGATGTAAGTAAAGATGAAGCGGCTGCAATGGCAGTAACTTAATCCC 900
Db 70319 ATTTTGCATGATGTAAGTAAAGATGAAGCGGCTGCAATGGCAGTAACTTAATCCC 70260
QY 901 AGCACTCTGGAGACTGAAGTGGGAGATCACTTGAGCTCAGGAGTTCAAGCAGCCTG 960
Db 70259 AGCACTCTGGAGACTGAAGTGGGAGATCACTTGAGCTCAGGAGTTCAAGCAGCCTG 70200
QY 961 GGCACCTAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1001
Db 70199 GGCACCTAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 70159

RESULT 9
HSA160H22/c HSA160H22 165436 bp DNA linear PRI 01-DEC-2000
LOCUS Human DNA sequence from clone RP11-160H22 on chromosome
DEFINITION 1q23.2-24.3, complete sequence.
ACCESSION AL121983
VERSION AL121983.13 GI:10443470
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 165436)
Bagguley, C.
Direct Submission
Submitted (30-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Oct 1, 2000 this sequence version replaced gi:9998793.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi, EMBL; Swi, SWISSPROT; Tri, TREMBL; Wp, WORMPEP; Information
```

on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
Chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>  
RP11-160H22 is from the library RPCI-11.1 constructed at the  
Roswell Park Cancer Institute by the group of Pieter de Jong. For  
further details see <http://bacpac.med.buffalo.edu/>  
VECTOR: pBACe3.6  
This sequence has been finished according to sequence map criteria  
as follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated repeat sequence elements. Where the sequence is  
ambiguous, there is an annotation using the 'unsure' feature key.  
IMPORTANT: This sequence is not the entire insert of clone  
RP11-160H22. It may be shorter because we sequence overlapping  
sections only once, except for a 100 base overlap.  
The true right end of clone RP11-160H22 is at 165436 in this  
sequence. The true left end of clone RP11-413B16 is at 81534 in  
this sequence. The true right end of clone RP4-681J21 is at 129090  
in this sequence. The true left end of clone RP5-1198E17 is at 100  
in this sequence.

## FEATURES

```
source
1..165436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q23.2-24.3"
/clone="RP11-160H22"
/clone_lib="RPCI-11.1"
repeat_region
1..152
/note="AluX repeat: matches 7..157 of consensus"
repeat_region
163..1009
/note="L1MC3 repeat: matches 6619..7483 of consensus"
repeat_region
1110..1363
/note="L1MC3 repeat: matches 6322..6604 of consensus"
repeat_region
1417..2313
/note="L1MC2 repeat: matches 5179..6113 of consensus"
repeat_region
2314..2654
/note="AluJb repeat: matches 2..310 of consensus"
repeat_region
2665..2776
/note="AluJb repeat: matches 8..133 of consensus"
repeat_region
2807..2985
/note="L1MC2 repeat: matches 4984..5164 of consensus"
repeat_region
3006..3320
/note="AluJb8 repeat: matches 1..315 of consensus"
repeat_region
3341..3380
/note="L1M4 repeat: matches 4443..4479 of consensus"
repeat_region
3381..3673
/note="AluX repeat: matches 1..296 of consensus"
repeat_region
3674..3779
/note="L1M4 repeat: matches 4325..4443 of consensus"
repeat_region
3870..4196
/note="AluJb8 repeat: matches 1..304 of consensus"
repeat_region
4243..4533
/note="AluX repeat: matches 1..300 of consensus"
repeat_region
4589..5103
/note="L1M4 repeat: matches 3804..4328 of consensus"
repeat_region
5181..5263
/note="L1MB4 repeat: matches 6095..6183 of consensus"
repeat_region
5445..5742
/note="AluX repeat: matches 1..298 of consensus"
repeat_region
5985..5994
/note="15 copies 2 mer ac 100% conserved"
repeat_region
5995..6018
/note="12 copies 2 mer ag 95% conserved"
repeat_region
6708..6987
/note="AluSp repeat: matches 1..297 of consensus"
repeat_region
7033..7339
/note="AluX repeat: matches 1..307 of consensus"
repeat_region
7405..7707
```

repeat\_region /note="Alusx repeat: matches 1. .303 of consensus"  
7812. .8120  
/note="AluJo repeat: matches 1. .311 of consensus"  
repeat\_region 8547. .8857  
/note="Alusx repeat: matches 1. .311 of consensus"  
repeat\_region 9267. .9520  
/note="Alusq repeat: matches 1. .254 of consensus"  
repeat\_region 10244. .10539  
/note="AluJo repeat: matches 2. .303 of consensus"  
misc\_feature complement(10251. .10663)  
/note="match: GSS: Em:AQ056360"  
11287. .11622  
/note="match: GSS: Em:AQ087768"  
12105. .12400  
/note="AluJo repeat: matches 1. .294 of consensus"  
repeat\_region 12406. .12701  
/note="Alusp repeat: matches 5. .311 of consensus"  
misc\_feature complement(13018. .13405)  
/note="match: GSS: Em:AQ470631"  
13041. .13258  
/note="WIR repeat: matches 11. .259 of consensus"  
repeat\_region 13309. .13442  
/note="L1MB8 repeat: matches 6034. .6171 of consensus"  
13685. .13816  
/note="WIR repeat: matches 47. .177 of consensus"  
14739. .15025  
/note="Alusx repeat: matches 1. .303 of consensus"  
16161. .16401  
/note="L2 repeat: matches 1594. .1811 of consensus"  
16402. .16677  
/note="Alusq repeat: matches 2. .278 of consensus"  
16582. .16914  
/note="match: STS: Em:HS278YE5"  
16678. .16885  
/note="L2 repeat: matches 1811. .2085 of consensus"  
16958. .17254  
/note="AluY repeat: matches 2. .297 of consensus"  
17394. .17708  
/note="Aluub repeat: matches 1. .307 of consensus"  
17819. .18069  
/note="Alusq repeat: matches 1. .239 of consensus"  
18343. .18456  
/note="WLRH repeat: matches 96. .211 of consensus"  
18522. .18710  
/note="WLRH repeat: matches 347. .535 of consensus"  
19219. .19529  
/note="Alusq repeat: matches 1. .311 of consensus"  
19681. .19751  
/note="WIR repeat: matches 124. .197 of consensus"  
20184. .20402  
/note="L1MB3A repeat: matches 5926. .6152 of consensus"  
20403. .20698  
/note="AluY repeat: matches 1. .295 of consensus"  
20699. .20754  
/note="L1MB3A repeat: matches 5872. .5926 of consensus"  
20757. .21052  
/note="AluJo repeat: matches 1. .302 of consensus"  
21990. .22293  
/note="Alusx repeat: matches 1. .304 of consensus"  
22631. .22699  
/note="L2 repeat: matches 2471. .2533 of consensus"  
22852. .23141  
/note="Alusx repeat: matches 1. .291 of consensus"  
23167. .23290  
/note="WIR repeat: matches 112. .249 of consensus"  
23452. .23493  
/note="L1MA2 repeat: matches 6267. .6308 of consensus"  
23575. .23682  
/note="L2 repeat: matches 2630. .2748 of consensus"  
23707. .23969  
/note="AluJb repeat: matches 40. .301 of consensus"  
25226. .25536  
/note="Aluub repeat: matches 1. .293 of consensus"

repeat\_region 25797. .26107  
/note="AluY repeat: matches 1. .309 of consensus"  
26292. .26357  
/note="33 copies 2 mer aa 68% conserved"  
26594. .26706  
/note="MER3 repeat: matches 33. .146 of consensus"  
26750. .27027  
/note="AlusC repeat: matches 1. .285 of consensus"  
27028. .27088  
/note="L1MC2 repeat: matches 5788. .5852 of consensus"  
27092. .27182  
/note="AluJ/FLAM repeat: matches 2. .91 of consensus"  
27217. .27673  
/note="L1MC2 repeat: matches 5847. .6321 of consensus"  
27674. .27734  
/note="MER3 repeat: matches 147. .209 of consensus"  
27695. .27745  
/note="MER3 repeat: matches 41. .89 of consensus"  
28664. .28959  
/note="Alusg1 repeat: matches 1. .292 of consensus"  
28986. .29101  
/note="Alusq/x repeat: matches 9. .124 of consensus"  
29214. .29504  
/note="Alusx repeat: matches 1. .291 of consensus"  
29721. .29852  
/note="FLAM C repeat: matches 2. .133 of consensus"  
30810. .31027  
/note="L1MB5 repeat: matches 5944. .6176 of consensus"  
  
Query Match 28.9%; Score 289.4; DB 9; Length 165436;  
Best Local Similarity 88.5%; Pred. No. 2.4e-63;  
Matches 314; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
  
Qy 1 AACAGAGAAAGCACATCTCGCGGGCGTGGTGGCTACACCTGTATATCCAACTTT 60  
Db 162103 AAAAAAAAAAAAAAAAAAGCGCGCGTGGCTCACCTGTATATCCAACTTT 162044  
  
Qy 61 GGGAGCGGCGGGTGGATCAGTCTAGGTACAGGATGGAGACCATCTGGCTGACACGG 120  
Db 162043 GGGAGCGGCGGGTGGATCAGGTCAGGAGATCGAGACCATCTGGCTAACACGG 161984  
  
Qy 121 TGAACCTCTCTCTACTTAAACACAAAAATAGCGGGCGTGGTGGAGCGGCGCTGT 180  
Db 161983 TGAACCTCTCTCTACTTAAACACAAAAATAGCGGGCGTGGTGGAGCGGCGCTGT 161924  
  
Qy 181 AGTCCAGTACTCGGAGGCTGAGGAGGAGAAATGCTTGAACCCGGAGCGCGAACTT 240  
Db 161923 AGTCCAGTACTCGGAGGCTGAGGAGGAGAAATGCTTGAACCCGGAGCGGAGCTT 161864  
  
Qy 241 GCAGTGAGCGAGTTGCACCATCGACTCGAGCTGGGCAACACAGTGAGACTCCGCT 300  
Db 161863 GCAGTGAGCGAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGAGACTCCGCT 161804  
  
Qy 301 CAAAAAAGAAAAAGAAAAAATCACTCTCATTCAGTGGTGGCACTT 355  
Db 161803 CAAAAAAGAAAAAAGAAAAAATTTTCCCATCTATAGTGGCTAT 161749  
  
RESULT 10  
AL359177  
LOCUS AL359177 230141 bp DNA linear HTG 10-JUL-2001  
DEFINITION Homo sapiens chromosome 1 clone RP11-413B16, 16 unordered pieces.  
ACCESSION AL359177  
VERSION AL359177.3 GI:12331067  
KEYWORDS HTG; HTGS PHASE1; HTGS\_CANCELLED.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Burton, J.  
TITLE Direct Submission  
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clome  
requests: clonerequest@sanger.ac.uk  
On Jan 22, 2001 this sequence version replaced gi:9797929.

## COMMENT

----- Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: BA413816  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Sequencing vector: pGAP4; 108752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 127350 bases at least Q40  
Consensus quality: 132189 bases at least Q30  
Consensus quality: 134528 bases at least Q20  
Insert size: 228641; sum-of-contigs  
Quality coverage: 152695; 8.1% error; agarose-fp  
Quality coverage: 2.04x in Q20 bases; sum-of-contigs Quality  
coverage: 3.11x in Q20 bases; agarose-fp  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 3155: contig of 3155 bp in length  
\* 3156 3255: gap of 100 bp  
\* 3256 10055: contig of 6800 bp in length  
\* 10056 10155: gap of 100 bp  
\* 10156 16833: contig of 6678 bp in length  
\* 16834 16933: gap of 100 bp  
\* 16934 36991: contig of 20058 bp in length  
\* 36992 37091: gap of 100 bp  
\* 37092 46051: contig of 8960 bp in length  
\* 46052 46151: gap of 100 bp  
\* 46152 58091: contig of 11940 bp in length  
\* 58092 58191: gap of 100 bp  
\* 58192 64171: contig of 5980 bp in length  
\* 64172 70251: contig of 5980 bp in length  
\* 70252 70351: gap of 100 bp  
\* 70352 73351: contig of 3000 bp in length  
\* 73352 73451: gap of 100 bp  
\* 73452 79431: contig of 5980 bp in length  
\* 79432 85511: gap of 100 bp  
\* 85512 85611: gap of 100 bp  
\* 85612 100531: contig of 14920 bp in length  
\* 100532 100631: gap of 100 bp  
\* 100632 106611: contig of 5980 bp in length  
\* 106612 106711: gap of 100 bp  
\* 106712 12459: contig of 18748 bp in length  
\* 124560 125559: gap of 100 bp  
\* 125560 134519: contig of 8960 bp in length  
\* 134520 134619: gap of 100 bp  
\* 134620 230141: contig of 95522 bp in length.

FEATURES  
source

1..230141  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP11-413B16"  
/clone\_lib="RPC1-11.2"  
1..3155  
/note="assembly\_fragment:00110"  
3256..10055  
/note="assembly\_fragment:00856"

misc\_feature

misc\_feature

misc\_feature 10156..16833  
/note="assembly\_fragment:01004"  
misc\_feature 16934..36991  
/note="assembly\_fragment:01205"  
misc\_feature 37092..46051  
/note="assembly\_fragment:01284"  
misc\_feature 46152..58091  
/note="assembly\_fragment:01287"  
misc\_feature 58192..64171  
/note="assembly\_fragment:01291"  
misc\_feature 64272..70251  
/note="assembly\_fragment:01293"  
misc\_feature 70352..73351  
/note="assembly\_fragment:01295"  
misc\_feature 73452..79431  
/note="assembly\_fragment:01300"  
misc\_feature 79532..85511  
/note="assembly\_fragment:01302"  
misc\_feature 85612..100531  
/note="assembly\_fragment:01305"  
misc\_feature 100632..106611  
/note="assembly\_fragment:01306"  
misc\_feature 106712..125459  
/note="assembly\_fragment:01310"  
misc\_feature 125560..134519  
/note="assembly\_fragment:01312"  
misc\_feature 134620..230141  
/note="assembly\_fragment:01255  
clone\_end:17  
vector\_side:right"

## ORIGIN

Query Match 28.9%; Score 289.4; DB 2; Length 230141;  
Best Local Similarity 88.5%; Pred. No. 2.5e-63;  
Matches 314; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
Qy 1 AACAGAGAAAGCACATCTCGCGCGCGGTGGTGCCTACACACCTGTATATCCACACTTT 60  
Db 149504 AAAAAAAAAAAAAAAAAAGCGCGCGGTGGTGCCTACACACCTGTATATCCACACTTT 149563  
Qy 61 GGGAGCGCGAGCGCGGTGGATCACTAGGTGAGGAGATGAGACCATCTGGCTGACACGG 120  
Db 149564 GGGAGCGCGAGCGCGGTGGATCACTAGGTGAGGAGATGAGACCATCTGGCTAACACGG 149623  
Qy 121 TGAACCCCTCTCTACTTAAACACAAAAATTTAGCCGGCGGTGGTGCAGCGCCTGT 180  
Db 149624 TGAACCCCTCTCTACTTAAACATACAAAAATTTAGCCGGCGGTGGTGCAGCGCCTGT 149683  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGCGCTGAACCCGGGAGCGGAACTT 240  
Db 149684 AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGCGCTGAACCCGGGAGCGGAGCTT 149743  
Qy 241 GCAGTGAGCGGAGTTGCACACTGCACTCAGCCTGGGCAACACAGTGAGACTCCGCT 300  
Db 149744 GCAGTGAGCGGAGTTGCGCACTCGCACTCAGCCTGGGCAACACAGTGAGACTCCGCT 149803  
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATCACTCTCATCAAGTGGTGCATTT 355  
Db 149804 CAAAAAAGAAAAAGAAAAAGAAATTTCTCCCATTCATATAGGTGGCTAT 149858

RESULT 11  
ALL138703/c

LOCUS ALL138703 129866 bp DNA linear PRI 21-APR-2001  
DEFINITION Human DNA sequence from clone RP11-413M19 on chromosome 13,  
complete sequence.  
ACCESSION AL138703  
VERSION AL138703.10 GI:13785012  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.





repeat\_region /note="MIR repeat: matches 73. .144 of consensus" 37119. .37244

repeat\_region /note="MERSA repeat: matches 73. .189 of consensus" 37713. .37765

repeat\_region /note="MSTD repeat: matches 1. .53 of consensus" 37766. .38064

repeat\_region /note="AluSq repeat: matches 1. .298 of consensus" 38065. .38409

repeat\_region /note="MSTD repeat: matches 53. .392 of consensus" 38512. .38630

repeat\_region /note="MIR repeat: matches 36. .155 of consensus" 38651. .38776

repeat\_region /note="L2 repeat: matches 1354. .1486 of consensus" 39209. .39304

repeat\_region /note="L2 repeat: matches 1818. .1924 of consensus" 40099. .40275

repeat\_region /note="MERSA repeat: matches 1. .168 of consensus" 41200. .41267

repeat\_region /note="34 copies 2 mer ac 80% conserved" 42055. .42551

repeat\_region /note="MLT2CA repeat: matches 1. .503 of consensus" 43484. .43529

repeat\_region /note="23 copies 2 mer ca 97% conserved" 45011. .45184

repeat\_region /note="MIR repeat: matches 15. .188 of consensus" 45185. .45740

repeat\_region /note="MLT2D repeat: matches 1. .551 of consensus" 45741. .45815

repeat\_region /note="MIR repeat: matches 188. .262 of consensus" 46651. .46956

repeat\_region /note="MER2 repeat: matches 1. .345 of consensus" 47341. .47633

repeat\_region /note="AluY repeat: matches 1. .293 of consensus" 47641. .48099

repeat\_region /note="MLT1D repeat: matches 1. .505 of consensus" 48139. .48307

repeat\_region /note="AluCb repeat: matches 130. .301 of consensus" 48373. .48498

repeat\_region /note="AluCb repeat: matches 1. .127 of consensus" 49476. .49630

repeat\_region /note="MIR repeat: matches 22. .190 of consensus" 50545. .50877

repeat\_region /note="L2 repeat: matches 1782. .2161 of consensus" 50946. .51187

Query Match 28.6%; Score 286; DB 9; Length 129866;  
Best Local Similarity 84.3%; Pred. No. 1.8e-62;  
Matches 322; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 1 AACAGAGAGACATCGCGCGGGCGTGGCTCACACCTGTATCCCAACTTT 60

Db 3209 AATAAGATAAACCATTCGCGCGGGCGCGGTGCTACCGCTGTATCCAGCACTTT 3150

Qy 61 GGGAGCCGAGCGGGTGGATCACTAGTCTAGGAGATCGACACCATCTGCTGACACGG 120

Db 3149 GGGAGCCGAGCGGGCGGATCACAGGTGAGGATCGAGACCATCTGCTGTAACCG 3090

Qy 121 TGAACCTCTCTACTATAAAACACAAAAAATTAGCCGGCGGTGGTGGAGGCGCTGT 180

Db 3089 TGAACCTCTCTACTATAAAATACAAAAAATTAGCCGGCGGTGGTGGAGGCGCTGT 3030

Qy 181 AGTCCAGCTACTCGGGAGGCTGAGCGAGGAGAAATGGCTCAACCCGGAGCGGAACCTT 240

Db 3029 AGTCCAGCTACTCGGGAGGCTGAGCGAGGAGAAATGGCTCAACCCGGAGCGGAACCTT 2970

Qy 241 GCAGTGAGCCGAGGTGACCACTGCACTCCAGCTGGGCAACACAGTGAGATCCGCTCT 300

Db 2969 GCAGTGAGCCGAGATCGCACCACTGCACTCCAGCTGGGCAACAGTGAGATCCGCTCT 2910

Qy 301 CAAAAAAG 360

Db 2909 CAAAAAAG 2850

Qy 361 TATTAGCCTTTCTGTAGCAA 382

Db 2849 TAAATAACATATCTTATAAAA 2828

RESULT 12

AL359177/c

LOCUS

DEFINITION Homo sapiens chromosome 1 clone RP11-413B16, 16 unordered pieces.

ACCESSION AL359177

VERSION AL359177.3 GI:12331067

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_CANCELLED.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Burton, J.

Direct Submission

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonrequest@sanger.ac.uk

On Jan 22, 2001 this sequence version replaced gi:9797929.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BA413B16

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 127350 bases at least Q40

Consensus quality: 132189 bases at least Q30

Consensus quality: 134528 bases at least Q20

Insert size: 228641; sum-of-contigs

Insert size: 152695; 8.1% error; agarose-fp

Quality coverage: 2.04x in Q20 bases; sum-of-contigs Quality coverage: 3.11x in Q20 bases; agarose-fp

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 16 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 3155: contig of 3155 bp in length

\* 3156 3255: gap of 100 bp

\* 3256 10055: contig of 800 bp in length

\* 10156 10155: gap of 100 bp

\* 16833 16933: contig of 6678 bp in length

\* 16934 16933: gap of 100 bp

\* 36991 37091: contig of 20058 bp in length

\* 36992 37091: gap of 100 bp

\* 37092 46051: contig of 8960 bp in length

\* 46052 46151: gap of 100 bp

\* 46152 58091: contig of 11940 bp in length

\* 58092 58191: gap of 100 bp

\* 58192 64171: contig of 5980 bp in length

\* 64172 64271: gap of 100 bp

\* 64272 70251: contig of 5980 bp in length

\* 70252 70351: gap of 100 bp

\* 70352 73451: contig of 3000 bp in length

\* 73452 73451: gap of 100 bp

\* 73452 79431: contig of 5980 bp in length

\* 79432 79531: gap of 100 bp

\* 79532 85511: contig of 5980 bp in length

\* 85512 85611: gap of 100 bp

\* 85612 100531: contig of 14920 bp in length



restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Oseegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.U. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-60F10, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-598C21; actual end is at base position 194197 of RP11-598C21.

Data from AC079249 was used to finish this clone, AC064871. In AC06487, the region from 90272 to 90395 is only represented by data from AC079249. Polymorphisms exist between RP11-598C21 and RP11-60F10.

#### FEATURES

source

Location/Qualifiers

1..194197

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-598C21"

/clone\_lib="RPCI-11"

70..328

/rpt\_family="MER1\_type"

repeat\_region

329..471

/rpt\_family="MIR"

repeat\_region

490..773

/rpt\_family="Alu"

misc\_feature

893..1028

/note="match to EST BF991908 (NID:g12398231)"

misc\_feature

893..1028

/note="match to EST BG120186 (NID:g12613708)"

misc\_feature

893..1028

/note="match to EST BG426467 (NID:g13332973)"

misc\_feature

893..1028

/note="match to EST BI093191 (NID:g14511521)"

misc\_feature

893..1028

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

misc\_feature

893..1028

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

misc\_feature

893..1019

/note="match to EST N44074 (NID:g1182602) yy30h01.r1"

misc\_feature

916..1028

/note="match to EST N4074 (NID:g1182602) yy30h01.r1"

misc\_feature

916..1028

/note="match to EST N4074 (NID:g1182602) yy30h01.r1"

misc\_feature

916..1028

/note="match to EST N4074 (NID:g1182602) yy30h01.r1"

misc\_feature

1892..2031

/note="match to EST BG426467 (NID:g13332973)"

misc\_feature

1892..2031

/note="match to EST BI093191 (NID:g14511521)"

misc\_feature

1892..2031

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

misc\_feature

1892..2031

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

misc\_feature

1892..2031

/note="similar to Mus musculus EST BF140223 (NID:g10979263)"

misc\_feature

1892..1957

/note="match to EST BF991908 (NID:g12398231)"

misc\_feature

1900..2031

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

misc\_feature

1971..2031

/note="match to EST AU125389 (NID:g10950105)"

repeat\_region

2576..2599

/rpt\_family="(TTTTC)n"

misc\_feature

2593..3011

/note="match to EST AA179659 (NID:g1761873) zp54f10.s1"

repeat\_region

3023..3128

/rpt\_family="L2"

repeat\_region

3287..3539

/rpt\_family="MER1\_type"

repeat\_region

3558..3597

/rpt\_family="L2"

repeat\_region

3757..3815

/rpt\_family="L1"

repeat\_region

6162..6191

/rpt\_family="(TAAA)n"

misc\_feature

6594..6713

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

misc\_feature

6594..6713

/note="match to EST AU125389 (NID:g10950105)"

misc\_feature

6594..6713

/note="match to EST AU133346 (NID:g10993885)"

misc\_feature

6594..6713

/note="match to EST BG120186 (NID:g12613708)"

misc\_feature

6594..6713

/note="match to EST BG426467 (NID:g13332973)"

misc\_feature

6594..6713

/note="match to EST BI093191 (NID:g14511521)"

misc\_feature

6594..6713

/note="similar to Homo sapiens EST BF242272 (NID:g1156199)"

misc\_feature

6594..6614

/note="match to EST N31469 (NID:g1151868) yx55g03.r1"

repeat\_region

7931..8013

/rpt\_family="L2"

repeat\_region

8280..8370

/rpt\_family="ERVL"

repeat\_region

8471..8516

/rpt\_family="AT-rich"

misc\_feature

9107..9243

/note="match to EST AW375848 (NID:g6880502)"

misc\_feature

9111..9243

/note="match to EST AI625215 (NID:g4650146) ts24b05.x1"

misc\_feature

9111..9243

/note="match to EST AU125389 (NID:g10950105)"

misc\_feature

9111..9243

Query Match 28.5%; Score 285; DB 9; Length 194197;  
Best Local Similarity 87.4%; Pred.No.3.4e-62;  
Matches 312; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 2 ACAGAGAAAGACATCTCGCCGGCGGTGGTGGCTACACCTGTATCCCAACACTTTG 61  
Db 15529 AAGAGATAAAAGATATCGCCGGCGGTGGTGGCTACACCTGTATCCCAACACTTTG 15588

Qy	62	GGAGGCCGAGCGCGGTGGATCACTAGCTCAGGAGATGGAGACCATCTCTGGCTGACACGGT	121
Db	15889	GGAGGCCGAGCGCGCGGATCACAGAGTCAAGAGATCGAGAGCATCCCGGCTAAAAACGGT	15648
Qy	122	GAAAACCTGCTCTACTATAAAAAACAATAAAATTTAGCCGGCGCTGGTGGCAGCGCGCTGTA	181
Db	15649	GAACCCCGTCTCTACTATAAAATACAAAAATTTAGCCGGCGCTAGTGGCGGCGCTGTA	15708
Qy	182	GTCCCGACTTACTCGGAGGCTGAGCGCAGGAGAAATGGCTTGAAACCGCGGAGCGCGAACTTG	241
Db	15709	GTCCCGAGCTACTTTGGAGGCTGAGCGCAGGAGAAATGGCGTGAACCCGGGAGCGCGAGCTTG	15768
Qy	242	CAGTGAAGCCAGGTTGCACCACTGCACCTTCAGCTCGGCAACACAGTGAAGACTCCGCTCTC	301
Db	15769	CAGTGAAGCCAGATCCCACTGCTCACTTCAGCTTGGGCGACAGAGCGAGACTCCGCTCTC	15828
Qy	302	AAAAAAAAAAAAAGAAAAAGAAATCACATCTCATTTCAAGTGGTGGCATTTAAA	358
Db	15829	AAAAAAAAAAAAAAAGAAAAAGATCTCAATTTCAATGAATTCGAAGAAATTGATA	15885

RESULT 14	
AL139231/c	
LOCUS	98868 bp DNA linear PRI 01-MAY-2001
DEFINITION	Human DNA sequence from clone RP11-439H9 on chromosome 6p21.33-22.3, complete sequence.
ACCESSION	AL139231
VERSION	AL139231.13 GI:13938786
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 98868)

**AUTHORS** Babbage, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (01-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENTS

On May 3, 2001 this sequence version replaced gi:13366275.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one W13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em1, EMBL; Sw:,  
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP  
database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/chr6>  
 RP11-439H9 is from the library RPC1-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>

vector: pBES.6  
 IMPORTANT: This sequence is not the entire insert of clone RP11-439H9. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-289W23 is at 98769 in this sequence.

FEATURES	Location/Qualifiers
source	1. .98868

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="p21.33-22.3"
/clone="RP11-439H9"
/clone_lib="RPC1-11.2"
5. 91
/note="L1M4 repeat: matches 2671. .2762 of consensus"
122. .201
/note="20 copies 4 mer tata 70% conserved"
227. .414
/note="L2 repeat: matches 2553. .2750 of consensus"
651. .1047
/note="THR1C repeat: matches 1. .371 of consensus"
2395. .2438
/note="22 copies 2 mer tt 79% conserved"
3036. .3910
/note="L2R12 repeat: matches 40. .2438 of consensus"
4467. .4757
/note="AluY repeat: matches 1. .291 of consensus"
4861. .4991
/note="MIR repeat: matches 124. .252 of consensus"
7740. .8050
/note="AluY repeat: matches 1. .311 of consensus"
8078. .8605
/note="MER61E repeat: matches 1. .525 of consensus"
10291. .10781
/note="MLR1D repeat: matches 1. .503 of consensus"
10818. .10849
/note="16 copies 2 mer ca 84% conserved"
13926. .14018
/note="MIR repeat: matches 53. .154 of consensus"
15593. .15739
/note="L1WA3 repeat: matches 6128. .6276 of consensus"
17234. .17505
/note="L2 repeat: matches 59. .322 of consensus"
19404. .19453
/note="25 copies 2 mer aa 80% conserved"
19788. .20190
/note="L2 repeat: matches 2346. .2750 of consensus"
20191. .20593
/note="MLR1A2 repeat: matches 1. .374 of consensus"
20594. .21281
/note="L2 repeat: matches 1560. .2346 of consensus"
21345. .21388
/note="11 copies 4 mer caca 93% conserved"
22058. .22374
/note="L2 repeat: matches 2358. .2657 of consensus"
22413. .22797
/note="THE1C repeat: matches 1. .371 of consensus"
22986. .23437
/note="L2 repeat: matches 2184. .2749 of consensus"
23516. .23937
/note="L2 repeat: matches 2277. .2735 of consensus"
25511. .25731
/note="MIR repeat: matches 33. .259 of consensus"
26900. .27030
/note="FLAM A repeat: matches 1. .131 of consensus"
28853. .28922
/note="L2 repeat: matches 2671. .2741 of consensus"
29632. .29693
/note="MLR1I repeat: matches 69. .132 of consensus"
29920. .30249
/note="AluY repeat: matches 1. .311 of consensus"
30508. .30725
/note="MIR repeat: matches 8. .236 of consensus"
30741. .31180
/note="L1PA15 repeat: matches 5704. .6145 of consensus"
31447. .31576
/note="FLAM C repeat: matches 1. .131 of consensus"
32749. .32864
/note="MLR1E repeat: matches 105. .225 of consensus"

```



REFERENCE  
 AUTHORS University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 TITLE 5 (bases 1 to 143947)  
 JOURNAL Waterston,R.  
 Direct Submission  
 Submitted (09-JAN-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT 6 (bases 1 to 143947)  
 Waterston,R.  
 Direct Submission  
 Submitted (01-MAR-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Jul 3, 2001 this sequence version replaced gi:13518276.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
 ----- Summary Statistics  
 -----  
 Center project name: H\_NH0566P21  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>  
 VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-804M4; the clone sequenced to the right is RP11-458D8, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-566P21; actual end is at base position 31704 of RP11-458D8.

Polymorphisms have been identified between AC073284 and AC012462.

#### FEATURES source

1. 143947  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-566P21"  
 /clone\_lib="RPCI-11"  
 misc\_feature 1..119  
 /note="match to EST AW364727 (NID:g6869481)"  
 misc\_feature 1..113  
 /note="match to EST A1752005 (NID:g5130269) cni13b05.xl"  
 misc\_feature 1..113  
 /note="match to EST AU117358 (NID:g10932320)"

misc\_feature 1..113  
 /note="match to EST AU138563 (NID:g11000084)"  
 misc\_feature 1..113  
 /note="match to EST AV752678 (NID:g10910526)"  
 misc\_feature 1..113  
 /note="match to EST AW068021 (NID:g6023019) cn20a08.xl"  
 misc\_feature 1..113  
 /note="match to EST AW608520 (NID:g7313261)"  
 misc\_feature 1..113  
 /note="match to EST BF751233 (NID:g12077909)"  
 misc\_feature 1..113  
 /note="match to EST BG163577 (NID:g12670280)"  
 misc\_feature 1..113  
 /note="match to EST BG260847 (NID:g12770663)"  
 misc\_feature 1..113  
 /note="similar to Bos taurus EST AV614913 (NID:g9750583)"  
 misc\_feature 1..113  
 /note="similar to EST BG897400 (NID:g14307641)"  
 misc\_feature 1..113  
 /note="similar to EST BG924585 (NID:g14319108)"  
 misc\_feature 1..113  
 /note="similar to EST BG927121 (NID:g14321644)"  
 misc\_feature 1..113  
 /note="similar to Mus musculus EST AL364292 (NID:g9691680)"  
 misc\_feature 8..113  
 /note="match to EST AI907037 (NID:g6497564)"  
 misc\_feature 45..113  
 /note="match to EST AU119133 (NID:g10934368)"  
 misc\_feature 57..113  
 /note="match to EST AW847780 (NID:g7943297)"  
 misc\_feature 59..113  
 /note="match to EST AU137394 (NID:g10997933)"  
 misc\_feature 63..113  
 /note="match to EST AU118498 (NID:g10933555)"  
 misc\_feature 78..113  
 /note="match to EST BE934096 (NID:g10460172)"  
 misc\_feature 78..113  
 /note="similar to EST BG993547 (NID:g14397617)"  
 misc\_feature 84..113  
 /note="match to EST AW608477 (NID:g7313218)"  
 misc\_feature 93..113  
 /note="match to EST BF339928 (NID:g11286390)"  
 misc\_feature 173..177  
 /note="similar to Homo sapiens EST AW058318 (NID:g5933957) wx18h04.xl"  
 misc\_feature 255..282  
 /note="match to EST A1378776 (NID:g4186629) tc18h12.xl"  
 repeat\_region 261..304  
 /rpt\_family="(TTTA)n"  
 repeat\_region 284..599  
 /rpt\_family="Alu"  
 misc\_feature 290..307  
 /note="match to EST A1271477 (NID:g3890644) q119f12.xl"  
 repeat\_region 450..481  
 /rpt\_family="(T)n"  
 misc\_feature 465..941  
 /note="match to EST A1333596 (NID:g4070155) qq06h09.xl"  
 repeat\_region 622..644  
 /rpt\_family="AT-rich"  
 misc\_feature 925..1258  
 /note="match to EST H77493 (NID:g1055582) yu09b11.r1"  
 misc\_feature 969..1339  
 /note="match to EST AA953560 (NID:g3117707) om95g02.s1"  
 misc\_feature 1019..1462  
 /note="match to EST D63177 (NID:g966846)"  
 misc\_feature 1319..1500  
 /note="match to EST BF326027 (NID:g11296775)"  
 misc\_feature 1722  
 /note="match to EST AA852317 (NID:g2940910)"  
 misc\_feature 1864..1899  
 /note="match to EST AV652269 (NID:g9873283)"  
 misc\_feature 1869..1899



```
misc_feature /note="match to EST AV652315 (NID:g9873329)"
1926..2448
misc_feature /note="match to EST BE169125 (NID:g8631846)"
1951..2289
misc_feature /note="match to EST AV652269 (NID:g9873283)"
1951..2289
misc_feature /note="match to EST AV652315 (NID:g9873329)"
2036..2037
misc_feature /note="match to EST BE812565 (NID:g10244903)"
2140
misc_feature /note="match to EST AV653321 (NID:g9874335)"
2219..2347
misc_feature /note="match to EST BF749100 (NID:g12075776)"
2322..2323
misc_feature /note="match to EST AW371903 (NID:g6876557)"
2345..2510
misc_feature /note="match to EST BG115942 (NID:g12609448)"
2346..2515
misc_feature /note="match to EST BF751233 (NID:g12077909)"
2346..2511
misc_feature /note="match to EST AW68021 (NID:g6023019) cn20a08.xl"
2346..2510
misc_feature /note="match to EST AI907037 (NID:g6497564)"
2346..2510
misc_feature /note="match to EST AU117358 (NID:g10932320)"
28.3%; Score 283.4; DB 9; Length 143947;
Query Match Best Local Similarity 90.7%; Pred. No. 8.4e-62;
Matches 302; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Qy 1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATATCCCAACACTTT 60
Db 60950 AAAAGTTGAAGAAGAAAGATCGCCGGCGCGTGGCTCACGCTGTATATCCAGCACTTT 60891
Qy 61 GGGAGCGCGAGCGCGGTGGATCACTAGTTCAGGAGATGGAGACCATCTGGCTGACACGG 120
Db 60890 GCGAGGCCAGCGCGGTGGATCAAGGTCAGGAGATCGAGACCATCTGGCTAACACGG 60831
Qy 121 TGAACCCCTGTCTTACTATAAAACACAAAAATTAGCCGGCGGTGGTGGAGCGCCTGT 180
Db 60830 TGAACCCCGTCTTACTATAAAATACAAAAATTAGCCGGCGGTGGTGGCGGTGCCTGT 60771
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGGAGCGGAACTT 240
Db 60770 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGATGGCTGAACCCGGGAGCGAGACTT 60711
Qy 241 GCAGTGAGCCGAGTTGCACCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCT 300
Db 60710 GCAGTGAGCCGAGATTGGCCACTGCACTCCAGCTGGGTGACAGAGTGAGACTCCGTCT 60651
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATACA 333
Db 60650 CAAAAAAGAAAAAAGAAAAAGAAAAAGA 60618
```

Search completed: July 4, 2005, 09:55:37  
Job time : 2892.32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:52:30 ; Search time 352.271 Seconds

(without alignments)  
16821.327 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_8000\_9000

Perfect score: 1001

Sequence: 1 aacagaagaagcaccatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	11570	3	AA95905 Human KIK
2	1001	100.0	11570	12	AdK52482 Human kal
3	1001	100.0	11570	13	AdR72623 Human ren
4	1001	100.0	11570	13	AdR72875 Human ova
5	283.4	28.3	31749	4	AAK72959 Human imm
6	280	28.0	78925	3	AAK89888 Human FN
7	279.4	27.9	96599	9	AAK157703 Human pro
8	279.4	27.9	96599	9	AdA02672 Human PRL
9	279.4	27.9	96599	10	AdB72410 Human PRL
10	278.4	27.9	96599	10	AdE95920 Human PRL
11	278.8	27.9	77781	10	AdL15049 Human mel
12	278.8	27.9	110300	13	AdS36499 Human aut
13	278.2	27.8	3407	4	AAK90223 Human dig
14	278.2	27.8	3407	5	AAK39852 Genomic s
15	278.2	27.8	3407	9	AdB32812 Human nov
16	277.6	27.7	347	4	AAK81654 Human imm
17	277.6	27.7	5115	4	AAK83483 Human imm
18	277.6	27.7	49777	4	AAK75029 Human imm
19	277.6	27.7	49777	4	AAK76214 Human imm
20	277.2	27.7	31952	4	AAK89370 Human dig

21	276.8	27.7	114793	4	AA080215 Human gen
22	276.6	27.6	119501	12	AdI29095 Human MAR
23	276.4	27.6	11474	4	AA105461 Human rep
24	276.4	27.6	31474	4	AB198314 Human tes
25	276.4	27.6	204621	11	ACN44486 Human gen
26	276.2	27.6	100445	13	ABD33179 Human can
27	275.2	27.5	41104	6	AA036260 Human G-p
28	274.8	27.5	2328	4	AA103890 Human rep
29	274.6	27.4	1400	12	AdH70101 Human Vbe
30	274.6	27.4	2325	4	AA103891 Human rep
31	274.6	27.4	3608	4	AAK83192 Human imm
32	274.6	27.4	3608	4	AAK74891 Human imm
33	274.6	27.4	3608	4	AAK67271 Human imm
34	274.6	27.4	110000	12	ADH69807_1 Continnuation (2 of
35	274.6	27.4	177380	8	ACF62751 Cancer ba
36	274.6	27.4	177380	8	AdB20870 MRPI base
37	274.6	27.4	177380	10	AdB87959 Human UGT
38	274.6	27.4	177380	10	AdB96942 Human MDR
39	274.6	27.4	177380	10	AdB92133 Human MDR
40	274.6	27.4	267156	6	AB168560 Kidney ca
41	274.4	27.4	465	5	ABV35226 Human pro
42	274.2	27.4	14543	6	ABK15798 Human von
43	274	27.4	1992	4	AAK85653 Human imm
44	274	27.4	1995	4	AAK85652 Human imm
45	274	27.4	17181	4	AAK72508 Human imm

## ALIGNMENTS

## RESULT 1

AAA95905  
ID AAA95905 standard; DNA; 11570 BP.  
XX  
AC AAA95905;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KIK-L2 gene.  
XX  
KW Human; KIK-L1; KIK-L2; KIK-L3; KIK-L4; KIK-L5; KIK-L6;  
KW kallikrein-like protein; serine protease; cytostatic; cancer;  
KW prostrate cancer; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-CA000258.  
XX  
PR 11-MAR-1999; 99US-0124260P.  
PR 01-APR-1999; 99US-0127386P.  
PR 21-JUL-1999; 99US-0144919P.  
PA (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
PI Aak72959 Human imm  
PI Aac89888 Human FN  
PI Aal57703 Human pro  
PI Yousef GM, Diamandis EP;  
DR WPI; 2000-587440/55.  
DR P-PSDB; AAB21296.  
XX  
XX New kallikrein-like (KIK-L) proteins for diagnosing and treating KIK-L  
PT protein mediated disorders, especially cancer.  
XX  
XX Claim 1; Page 143-149; 184pp; English.  
XX  
XX The present sequence is the coding sequence of the human KIK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyse the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding



Db 8120 TGAACCCCTGTCTCTACTAAAAACACAAAAAATTAGCCGGCGTGTGGCAGGCGCCTGT 8179  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAAACCCGGAGGCGGAACCTT 240  
Db 8180 AGTCCAGCTACTCGGAGGCTGAGGAGGAGATGGCTGAAACCCGGAGGCGGAACCTT 8239  
Qy 241 GCAGTGAGCCGAGGTTGCACCACTGCCTCCAGGCTGGGCAACACAGTGAGACTCCGCTCT 300  
Db 8240 GCAGTGAGCCGAGGTTGCACCACTGCCTCCAGGCTGGGCAACACAGTGAGACTCCGCTCT 8299  
Qy 301 CAAAAAAGAAAAAGAAAAAGAAATCAATCTCATTCATCAAGTGGTGGCATTTAAAAAC 360  
Db 8300 CAAAAAAGAAAAAGAAAAAGAAATCAATCTCATTCATTCAGTGGTGGCATTTAAAAAC 8359  
Qy 361 TATTAGCCTTCTCTAGGCAAGCTTAGTACTCTGTTTTCAGACCTCAAGTGTGTTTTT 420  
Db 8360 TATTAGCCTTCTCTAGGCAAGCTTAGTACTCTGTTTTCAGACCTCAAGTGTGTTTTT 8419  
Qy 421 TTGTTTGTGTTTTTTCATACCGGTGTGTCTGGGTGGGCCACTAAAAAGCTTACAAGCAAG 480  
Db 8420 TTGTTTGTGTTTTTTCATACCGGTGTGTCTGGGTGGGCCACTAAAAAGCTTACAAGCAAG 8479  
Qy 481 AAATAATAACAACACTACAACACTAATAACCAATAGTATATAAATAATAGCATCTGGCTA 540  
Db 8480 AAATAATAACAACACTACAACACTAATAACCAATAGTATATAAATAATAGCATCTGGCTA 8539  
Qy 541 ATTGCTGACACTGTTTAAAGTGGTTCATGCTGCTCAGCTCATTAATCATCTACCTGTT 600  
Db 8540 ATTGCTGACACTGTTTAAAGTGGTTCATGCTGCTCAGCTCATTAATCATCTACCTGTT 8599  
Qy 601 ATTATGGCCCTATTATCAACAGGAGGCAAGGCTCAGAGCAGTAACTAAACAGGCTC 660  
Db 8600 ATTATGGCCCTATTATCAACAGGAGGCAAGGCTCAGAGCAGTAACTAAACAGGCTC 8659  
Qy 661 TCAAAAGAAAACTGCGAGAGATATTAATTTAAAAAATAATGAGAGAAATTAACCCACAA 720  
Db 8660 TCAAAAGAAAACTGCGAGAGATATTAATTTAAAAAATAATGAGAGAAATTAACCCACAA 8719  
Qy 721 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 780  
Db 8720 GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGCTTTGAAACAGTGTCTGCT 8779  
Qy 781 ACTGGAAAAAGGCAAGTCTTGGCTTCTTAATTAATGATACCAAGACTCTGTAATTCAT 840  
Db 8780 ACTGGAAAAAGGCAAGTCTTGGCTTCTTAATTAATGATACCAAGACTCTGTAATTCAT 8839  
Qy 841 ATTTTGCATGCATGTAAGTAAAGAAATGAAGCGGTGCAATGGCACATGCCAGTAATCCC 900  
Db 8840 ATTTTGCATGCATGTAAGTAAAGAAATGAAGCGGTGCAATGGCACATGCCAGTAATCCC 8899  
Qy 901 AGCACTCTGGGAGCTGAAGTGGGAGATCACTTGAAGCTCAGGAGTTCAAGCAGGCTG 960  
Db 8900 AGCACTCTGGGAGCTGAAGTGGGAGATCACTTGAAGCTCAGGAGTTCAAGCAGGCTG 8959  
Qy 961 GGCACCTAAAAATTAATAAATAATAAATACTAATGTTTTT 1001  
Db 8960 GGCACCTAAAAATTAATAAATAATAAATACTAATGTTTTT 9000

## RESULT 3

ADNR72623  
ID ADNR72623 standard; DNA; 11570 BP.

XX AC ADNR72623;

XX DT 02-DEC-2004 (first entry)

XX DE Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.

XX KW kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;

XX KW cancer metastasis; chemotherapy; human; serine protease;

XX KW chromosome 19q13.4; KLK5; ds; gene.

OS Homo sapiens.

XX Key Location/Qualifiers  
PH CDS 2221..11247

FT /\*tag= b  
FT /product= "Human renal cell carcinoma-related kallikrein  
FT 5 (hK5) protein"

FT exon 2221..2293

FT /\*tag= a

FT /number= 1

FT intron 2294..4761

FT /\*tag= c

FT /number= 1

FT exon 4762..5023

FT /\*tag= d

FT /number= 2

FT intron 5024..5762

FT /\*tag= e

FT /number= 2

FT exon 5763..6019

FT /\*tag= f

FT /number= 3

FT intron 6020..6104

FT /\*tag= g

FT /number= 3

FT exon 6105..6238

FT /\*tag= h

FT /number= 4

FT exon 6239..11091

FT /\*tag= i

FT /number= 4

FT exon 11092..11247

FT /\*tag= j

FT /number= 5

FT

FT WO2004077060-A2.

PN 10-SEP-2004.

XX 26-FEB-2004; 2004WO-CA000280.

XX 27-FEB-2003; 2003US-0451382P.

PR (MOUN ) MOUNT SINAI HOSPITAL.

PA Diamandis EP, Petraki CD;

XX WPI; 2004-662077/64.

XX P-PSDB; ADR72621.

XX Detecting kallikrein polypeptides/polynucleotides associated with renal

XX cell carcinoma in a patient, for diagnosing/treating the disease,

XX comprises detecting /identifying kallikrein polypeptides/polynucleotides

XX in a sample.

XX Example 1; SEQ ID NO 3; 53pp; English.

XX The invention relates to a novel method for detecting kallikrein

XX polypeptides, or the polynucleotides encoding them, associated with renal

XX cell carcinoma. The method comprises obtaining a sample from a patient

XX and detecting kallikrein polypeptides, or their encoding polynucleotides,

XX where the kallikrein polypeptides are selected from the group consisting

XX of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The

XX detected amounts of the kallikrein polypeptides are compared to standard

XX amounts. The molecules of the invention demonstrate cytostatic activity

XX whilst the methods and kit may be useful for detecting, characterising,

XX preventing and treating renal cell carcinoma. Furthermore, the methods

XX may be useful for evaluating the probability of the presence of malignant

XX or pre-malignant cells and for detecting and quantitating tumour growth

XX and cancer metastasis. Finally, the methods may be utilised to confirm

XX the absence or removal of all tumour tissue following surgery, cancer

XX chemotherapy and/or radiation therapy and to monitor cancer chemotherapy

XX and tumour reappearance. The current sequence is that of the human

CC	ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the									
CC	invention which encodes a secreted serine protease and is located at									
CC	chromosome 19q13.4.									
XX										
SQ	Sequence	11570 BP;	3034 A;	2562 C;	3325 G;	2649 T;	0 U;	0 Other;		
	Query Match	100.0%; Score 1001; DB 13; Length 11570;								
	Best Local Similarity	100.0%; Pred. No. 2.9e-239;								
	Matches 1001; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	1	AACAGAGAAAGACATCTCGCGCGGGCTGGTGGCTCACACCTGTATCCCAACACTTT	60							
DB	8000	AACAGAGAAAGACATCTCGCGCGGGCTGGTGGCTCACACCTGTATCCCAACACTTT	8059							
QY	61	GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTCGCTCACACGG	120							
DB	8060	GGGAGGCGGAGCGGGTGGATCACTAGTCAAGAGATGGAGACCATCTCGCTCACACGG	8119							
QY	121	TGAAACCCCTGTCCTACTTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCTGT	180							
DB	8120	TGAAACCCCTGTCCTACTTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCTGT	8179							
QY	181	AGTCCCAGCTACTCGGAGGCTGAGGCAGGAGATGGCTGAACCCGGGAGCGGAACTT	240							
DB	8180	AGTCCCAGCTACTCGGAGGCTGAGGCAGGAGATGGCTGAACCCGGGAGCGGAACTT	8239							
QY	241	GCAGTGAGCCGAGTTGCACCACTGCATCTCAGAGCTGGGCAACACAGTGAGACTCCGTCT	300							
DB	8240	GCAGTGAGCCGAGTTGCACCACTGCATCTCAGAGCTGGGCAACACAGTGAGACTCCGTCT	8299							
QY	301	CAAAAAAAGAAAAAGAAAAAATCAATCATCTCATCTCAAGTGGTGGCAATTTAAAAAC	360							
DB	8300	CAAAAAAAGAAAAAGAAAAAATCAATCATCTCATCTCAAGTGGTGGCAATTTAAAAAC	8359							
QY	361	TATTTAGCTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT	420							
DB	8360	TATTTAGCTTCTGTAGGCAAGTTAGTATCTTTTCCAGACCTCAAGGTGTTTTT	8419							
QY	421	TTGTTTGTGTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAAGCTACAGCAAG	480							
DB	8420	TTGTTTGTGTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAAGCTACAGCAAG	8479							
QY	481	AAATAATAACAACATCAACAATCTAATACCAATAGTATAAAAAATAATAGCATCTGGCTA	540							
DB	8480	AAATAATAACAACATCAACAATCTAATACCAATAGTATAAAAAATAATAGCATCTGGCTA	8539							
QY	541	ATTGCTGACACTGTTTAAAGTGGTTTCGATGCCTCAGCTCATTAATCAATTAACCTGTT	600							
DB	8540	ATTGCTGACACTGTTTAAAGTGGTTTCGATGCCTCAGCTCATTAATCAATTAACCTGTT	8599							
QY	601	ATTATTGGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCAGTTAACTAACAGGCTC	660							
DB	8600	ATTATTGGCCCTATTTTACAAACAGGAGCAAGGCTCAGAGCAGTTAACTAACAGGCTC	8659							
QY	661	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTAACCCCAA	720							
DB	8660	TCAAAAGAAACTCTGCAGAGATATTAAATTTAAAAAATAATGAGAGAAATTAACCCCAA	8719							
QY	721	GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGTTTGTGAAACAGTGTCTGCT	780							
DB	8720	GAAAGTTGAAATTTAGAGGTACAGCAGCTAAGCTTGTGTTTGTGAAACAGTGTCTGCT	8779							
QY	781	ACTGGGAAAAAGGCAAGCTTTGGCTTTCCTAATTAATTTGATACCAAGGACTCTGTAATTCAT	840							
DB	8780	ACTGGGAAAAAGGCAAGCTTTGGCTTTCCTAATTAATTTGATACCAAGGACTCTGTAATTCAT	8839							
QY	841	ATTTTGCATGATGTAAGTAAGAAATGAAGCGGGTGCAATGGCAGCATGCCAGTAATCCC	900							
DB	8840	ATTTTGCATGATGTAAGTAAGAAATGAAGCGGGTGCAATGGCAGCATGCCAGTAATCCC	8899							
QY	901	AGCACTCGGGAGACTGAAGTGGGAGATCATCTTGAGCTCAGGAGTTCAAGACGAGCTG	960							
DB	8900	AGCACTCGGGAGACTGAAGTGGGAGATCATCTTGAGCTCAGGAGTTCAAGACGAGCTG	8959							

Qy 961 GGCAACTAAAAATTAAAAATAAAAAATAAAATACTAAATGTTTTT 1001  
|||||  
Db 8960 GGCAACTAAAAATTAAAAATAAAAAATAAAATACTAAATGTTTTT 9000  
|||||

RESULT 4  
ADR72875

ID	ADR72875	standard; DNA; 11570 BP.
XX	ADR72875;	
XX	DT	
XX	02-DEC-2004	(first entry)
DE	Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.	
XX	kallikrein 5; tumour marker; ovarian cancer;	
KW	epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;	
KW	KLK5; ds; gene.	
XX	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	2221..11247
FT		/*tag= b
FT		/product= "Human ovarian cancer-related tumour marker
FT		kallikrein 5 (hk5) protein"
FT	exon	2221..2293
FT		/*tag= a
FT		/number= 1
FT	intron	2294..4761
FT		/*tag= c
FT		/number= 1
FT	exon	4762..5023
FT		/*tag= d
FT		/number= 2
FT	intron	5024..5762
FT		/*tag= e
FT		/number= 2
FT	exon	5763..6019
FT		/*tag= f
FT		/number= 3
FT	intron	6020..6104
FT		/*tag= g
FT		/number= 3
FT	exon	6105..6238
FT		/*tag= h
FT		/number= 4
FT	exon	6239..11091
FT		/*tag= i
FT		/number= 4
FT	exon	11092..11247
FT		/*tag= j
FT		/number= 5

WO2004075713-A2.

10-SEP-2004.

26-FEB-2004; 2004WO-CA000281.

26-FEB-2003; 2003US-0450406P.

(MOUN ) MOUNT SINAI HOSPITAL.

Diamandis EP;

WPI; 2004-661815/64.

P-PSDB; ADR72873.

Kallikrein markers detection method for detecting ovarian cancer in patient, involves detecting kallikrein markers and Cal25 in sample obtained from patient, and comparing detected amounts with standard

PT amounts.  
XX Example 2; SEQ ID NO 5; 102pp; English.  
PS  
XX  
XX  
The invention relates to a novel method for detecting a plurality of kallikrein markers associated with ovarian cancer. The method comprises obtaining a sample from a patient and detecting in the sample a plurality of kallikrein markers, and optionally carbohydrate antigen CA125, wherein the kallikrein markers are selected from the group consisting of kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and kallikrein 11. The detected amounts of the kallikrein markers are compared with standard amounts. The method of the invention may be useful for detecting kallikrein markers associated with ovarian cancer in a patient and thus for detecting ovarian cancer, particularly epithelial ovarian carcinoma. The current sequence is that of the human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention which encodes a secreted serine protease and is located at chromosome 19q13.4.  
XX  
XX  
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 13; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 2.9e-239;  
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACAGAGAACACATCTCGCGCGGCGTGTGGCTCACACCTGTATATCCACACTTT 60  
DB 8000 AACAGAGAACACATCTCGCGCGGCGTGTGGCTCACACCTGTATATCCACACTTT 8059  
QY 61 GGGAGCGGAGCGGTGATCAGTGTAGGATGGAGACCATCTCTGCTGACACGG 120  
DB 8060 GGGAGCGGAGCGGTGATCAGTGTAGGATGGAGACCATCTCTGCTGACACGG 8119  
QY 121 TGAACCCCTGTCTTACTATAAACAACAAAAATAGCCGGCGTGTGGCAGCGCCCTGT 180  
DB 8120 TGAACCCCTGTCTTACTATAAACAACAAAAATAGCCGGCGTGTGGCAGCGCCCTGT 8179  
QY 181 AGTCCAGCTACTCGGGAGGCTGAGGAGGAGGAGATGGCTGAAACCCGGAGGCGGAACTT 240  
DB 8180 AGTCCAGCTACTCGGGAGGCTGAGGAGGAGGAGATGGCTGAAACCCGGAGGCGGAACTT 8239  
QY 241 CGAGTGAGCCGAGGTTGACCACTGCATCTCAGCTGGGCAACAGTGAGCTCCGCT 300  
DB 8240 CGAGTGAGCCGAGGTTGACCACTGCATCTCAGCTGGGCAACAGTGAGCTCCGCT 8299  
QY 301 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGTGGCATTTAAAAAC 360  
DB 8300 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGTGGCATTTAAAAAC 8359  
QY 361 TATTAGCCTTCTCTAGGCAAGTTAGTATCTTCTTTTCCAGACCTCAAGGTGTTTTT 420  
DB 8360 TATTAGCCTTCTCTAGGCAAGTTAGTATCTTCTTTTCCAGACCTCAAGGTGTTTTT 8419  
QY 421 TTGTTTGTGTTTTCATACGGTGTGTGTCTGGGTGTCGCCACTAAAGCTACAGCAAG 480  
DB 8420 TTGTTTGTGTTTTCATACGGTGTGTGTCTGGGTGTCGCCACTAAAGCTACAGCAAG 8479  
QY 481 AAATAATAACAACACTACAACATACTAATAACCAATAGTATAAATAATAGCATCTGGCTA 540  
DB 8480 AAATAATAACAACACTACAACATACTAATAACCAATAGTATAAATAATAGCATCTGGCTA 8539  
QY 541 ATTGCTGACACTGTTTTTAAGTGGTTTGATGCTCAGCTCATTTACCTGTT 600  
DB 8540 ATTGCTGACACTGTTTTTAAGTGGTTTGATGCTCAGCTCATTTACCTGTT 8599  
QY 601 ATTATTGGCCCTATTTTTCAACAACAGGAGCCAGGCTCAGACGTTAACTAACAGCTC 660  
DB 8600 ATTATTGGCCCTATTTTTCAACAACAGGAGCCAGGCTCAGACGTTAACTAACAGCTC 8659  
QY 661 TCAAAAGAAACTCTCAGAGATATTAAATTTTAAAAATAATGAGAGAAATTAACCAAA 720  
DB 8660 TCAAAAGAAACTCTCAGAGATATTAAATTTTAAAAATAATGAGAGAAATTAACCAAA 8719

QY 721 GAAAGTTGAAATTTAGAGGTACAGGAGCTTAAGCTTGTGTTTGAACAGTGTCTGCT 780  
DB 8720 GAAAGTTGAAATTTAGAGGTACAGGAGCTTAAGCTTGTGTTTGAACAGTGTCTGCT 8779  
QY 781 ACTGGGAAAAAGGCAAGTCTTGGCTTTCCCTAATAATTGATACACAGGACTCTGTAATTCAT 840  
DB 8780 ACTGGGAAAAAGGCAAGTCTTGGCTTTCCCTAATAATTGATACACAGGACTCTGTAATTCAT 8839  
QY 841 ATTTTGCATGCATTAAGTAAGAAATGAAGCCGGTGCAATGGCACATGCCAGTAATCCC 900  
DB 8840 ATTTTGCATGCATTAAGTAAGAAATGAAGCCGGTGCAATGGCACATGCCAGTAATCCC 8899  
QY 901 AGCACTCTGGGAGACTGAAGTGGGAGGATCACTTGGCTCAGGAGTTCAAGCACCAGCTG 960  
DB 8900 AGCACTCTGGGAGACTGAAGTGGGAGGATCACTTGGCTCAGGAGTTCAAGCACCAGCTG 8959  
QY 961 GGCAACTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1001  
DB 8960 GGCAACTAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9000  
RESULT 5  
AAK72959  
ID AAK72959 standard; DNA; 31749 BP.  
XX AAK72959;  
AC AAK72959;  
XX  
DT 06-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27771.  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
KW Homo sapiens.  
OS  
XX WO200157182-A2.  
PN 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001354.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.

PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226686P.	PR	17-NOV-2000;	2000US-0249208P.
PR	22-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249246P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235834P.	XX		
PR	27-SEP-2000;	2000US-0235836P.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	29-SEP-2000;	2000US-0236327P.	XX	WPI; 2001-483426/52.	
PR	29-SEP-2000;	2000US-0236367P.	DR		
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	29-SEP-2000;	2000US-0236370P.	PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	02-OCT-2000;	2000US-0236802P.	XX	Disclosure; SEQ ID NO 27771; 3071pp + Sequence Listing; English.	
PR	02-OCT-2000;	2000US-0237037P.	PS		
PR	02-OCT-2000;	2000US-0237038P.	XX		
PR	02-OCT-2000;	2000US-0237039P.	CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	02-OCT-2000;	2000US-0237040P.	CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
PR	13-OCT-2000;	2000US-0239935P.	CC	activity, and can be used in gene therapy and vaccine production. (I)	
PR	13-OCT-2000;	2000US-0239937P.	CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
PR	20-OCT-2000;	2000US-0240960P.	CC	treatment of diseases associated with inappropriate (I) expression. For	
PR	20-OCT-2000;	2000US-0241221P.	CC	example, they may be used to treat disorders associated with decreased	
PR	20-OCT-2000;	2000US-0241785P.	CC	expression by rectifying mutations or deletions in a patient's genome	
PR	20-OCT-2000;	2000US-0241786P.	CC	that affect the activity of (I) by expressing inactive proteins or to	
PR	20-OCT-2000;	2000US-0241787P.	CC	supplement the patients own production of (I). Additionally, (I)	
PR	20-OCT-2000;	2000US-0241808P.	CC	polynucleotides may be used to produce the secreted (I), by inserting the	
PR	20-OCT-2000;	2000US-0241809P.	CC	nucleic acids into a host cell and culturing the cell to express the	
PR	20-OCT-2000;	2000US-0241826P.	CC	protein. (I) proteins and polynucleotides may be used to prevent,	
PR	01-NOV-2000;	2000US-0244617P.	CC	diagnose and treat immune/haematopoietic-related diseases, especially	
PR	08-NOV-2000;	2000US-0246474P.	CC	cancers and cancer metastases of haematopoietic-derived cells. AAK4703	
PR	08-NOV-2000;	2000US-0246475P.	CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
PR	08-NOV-2000;	2000US-0246476P.	CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
PR	08-NOV-2000;	2000US-0246477P.	CC	represent sequences used in the exemplification of the present invention	
PR	08-NOV-2000;	2000US-0246478P.	XX		
PR	08-NOV-2000;	2000US-0246523P.	SQ	Sequence 31749 BP; 9613 A; 6041 C; 6509 G; 9586 T; 0 U; 0 Other;	
PR	08-NOV-2000;	2000US-0246524P.			
PR	08-NOV-2000;	2000US-0246525P.			
PR	08-NOV-2000;	2000US-0246526P.			
PR	08-NOV-2000;	2000US-0246527P.			
PR	08-NOV-2000;	2000US-0246528P.			
PR	08-NOV-2000;	2000US-0246532P.			
PR	08-NOV-2000;	2000US-0246609P.			
PR	08-NOV-2000;	2000US-0246610P.			
PR	08-NOV-2000;	2000US-0246611P.			
PR	08-NOV-2000;	2000US-0246613P.			
<hr/>					
Qy		1 AACAGAGAAAGCACATCTCGCGCGGCGTGGTGGCTCACACCTGTATATCCAAACATTT 60			
Db		9594 AAAAGTTGAAGAAAAGATCGCGCGCGCGTGGCTCACACCTGTATATCCAGCACHTT 9653			
Qy		61 GGGAGGCGGAGCGGCGTGGATCACTAGGTCTAGGAGATGGAGACCATCTGCTGACACGG 120			
<hr/>					
Query Match		28.3%;	Score 283.4;	DB 4;	Length 31749;
Best Local Similarity		90.7%;	Pred. No. 5.1e-60;		
Matches 302;		Conservative 0;	Mismatches 31;	Indels 0;	Gaps 0;
<hr/>					





overproduction of which is associated with cancer as well as various autoimmune diseases. The invention may provide a means of inhibiting the activity of carcinoma-associated proteins (CAP) through use in antisense therapy or gene therapy, thus exhibiting cytostatic activity. The nucleic acid sequence of the invention may be useful for diagnosing carcinomas and also as probes to determine the chromosomal location of CA genes and the number of copies of the CA genes. They may also be useful in generating animal models of carcinomas, particularly breast cancer or as DNA vaccines. The CAP protein may be useful for screening for a bioactive agent capable of binding to CAP or capable of modulating the activity of CAP. Suitable cancers which may be diagnosed or screened include cancer of the oral cavity and pharynx, cancers of the digestive system, cancers of the respiratory system, mesothelioma, cancer of the bones and joints, and soft tissue, including heart, skin cancers, cancer of the urinary system, endocrine system, multiple myeloma, and leukaemia. The CAP may also be useful as a marker of carcinoma (including breast cancer) or lymphoma (such as Hodgkin's and non-Hodgkin's lymphoma). The present sequence is the genomic sequence of the human prolactin receptor (PRLR) gene, a CA nucleic acid sequence of the invention

PR 26-DEC-2001; 2001US-00035832.  
XX (SAGR-) SAGRES DISCOVERY.  
XX Morris DW;  
XX WPI; 2003-587068/55.  
DR New recombinant nucleic acid encoding carcinoma associated protein,  
PT useful for preparing compositions for treating carcinomas.  
XX Claim 1; SEQ ID NO 1190; 245pp; English.  
XX The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Query Match 27.9%; Score 279.4; DB 9; Length 96599;  
Best Local Similarity 90.6%; Pred. No. 7.4e-59;  
Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 GAAGAAAGCAGCATCTCGGCGGGCGTGGCTCACACCTGTAAATCCCAACACTTTGGGA 64  
DB GAAATAACATCTGGGCGGGCGGTGGCTCAGCGCTGTAAATCCAGCACCTTTGGGA 49103  
QY 65 GGCCGAGCGGGTGGATCACTAGGTCAGGAGATGAGACCATCTCTGGCTGACACGGTGAA 124  
DB GGCCGAGCGGGCGGATCAGGAGTTCAGGAGACCATCTCCGGCTAAACGGTGAA 49163  
QY 125 ACCCTGTCTCTACTAAAAACACAAAAATTAGCCGGGCGTGGTGCAGCGCCCTGTAGTC 184  
DB ACCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGCAGCGCCCTGTAGTC 49223  
QY 185 CCAGTACTTCGGAGGCTGAGCAGGAGATGGCTGAAACCCGGGAGCGGAACTTGCAG 244  
DB CCAGTACTTCGGAGGCTGAGCAGGAGATGGCTGAAACCCGGGAGCGGAGCTTGCAG 49283  
QY 245 TGAGCCGAGGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGCTCTCAA 304  
DB TGAGCCGAGATTGGCCCACTGCACTCCAGCCTGGGCGAGAGAGAGACTCCGCTCTCAA 49343  
QY 305 AAAAAAAGAAAGAAAGAAATCACA 333  
DB AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 49372

## RESULT 8

ADA02672  
ID ADA02672 standard; DNA; 96599 BP.

XX AC ADA02672;

XX DT 06-NOV-2003 (first entry)

XX DE Human PRLR carcinoma associated gene, SEQ ID NO:1190.

XX KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;  
XX KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;  
XX KW gene; ds.

XX OS Homo sapiens.

XX PN WO2003057146-A2.

XX PD 17-JUL-2003.

XX PF 26-DEC-2002; 2002WO-US041414.

XX

PR

26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

DR New recombinant nucleic acid encoding carcinoma associated protein,  
PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1190; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic acid sequences from mouse and human (ADA01482-ADA03094), and to recombinant carcinoma associated proteins (CAP) encoded by them. The invention also encompasses expression vectors and host cells comprising a CA nucleic acid, a polypeptide (especially an antibody) that specifically binds to the protein, and a biochip comprising CA nucleic acid or fragments thereof. The sequences of the invention were identified using oncogenic retroviruses, which insert into the genome of the host organism at random. Many of these do not carry transduced host oncogenes or pathogenic trans-acting viral genes, meaning that cancer incidence is a direct consequence of the effects of proviral integration into host protooncogenes. The CA nucleic acid sequences can be used to diagnose carcinoma (especially breast cancer, prostate cancer, lymphoma or leukaemia) or a propensity to carcinoma by determination of the sequence of a CA gene, or by determination of CA gene expression in particular tissues. CA nucleic acids, proteins and antibodies are also useful as therapeutic agents and in screening and evaluating drug candidates. The present sequence represents a specifically claimed human CA nucleic acid sequence of the invention. Note: The complete sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 96599 BP; 28791 A; 19399 C; 19931 G; 28478 T; 0 U; 0 Other;

Query Match 27.9%; Score 279.4; DB 9; Length 96599;

Best Local Similarity 90.6%; Pred. No. 7.4e-59;

Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 5 GAAGAAAGCAGCATCTCGGCGGGCGTGGCTCACACCTGTAAATCCCAACACTTTGGGA 64

DB GAAATAACATCTGGGCGGGCGGTGGCTCAGCGCTGTAAATCCAGCACCTTTGGGA 49103

QY 65 GGCCGAGCGGGTGGATCACTAGGTCAGGAGATGAGACCATCTCTGGCTGACACGGTGAA 124

DB GGCCGAGCGGGCGGATCAGGAGTTCAGGAGACCATCTCCGGCTAAACGGTGAA 49163

QY 125 ACCCTGTCTCTACTAAAAACACAAAAATTAGCCGGGCGTGGTGCAGCGCCCTGTAGTC 184

DB ACCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGCAGCGCCCTGTAGTC 49223

QY 185 CCAGTACTTCGGAGGCTGAGCAGGAGATGGCTGAAACCCGGGAGCGGAACTTGCAG 244

DB CCAGTACTTCGGAGGCTGAGCAGGAGATGGCTGAAACCCGGGAGCGGAGCTTGCAG 49283

QY 245 TGAGCCGAGGTTGCACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGCTCTCAA 304

DB TGAGCCGAGATTGGCCCACTGCACTCCAGCCTGGGCGAGAGAGACTCCGCTCTCAA 49343

QY 305 AAAAAAAGAAAGAAAGAAAGAAATCACA 333

DB AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 49372

## RESULT 9

ADB72410

ID ADB72410 standard; DNA; 96599 BP.

XX AC ADB72410;

XX

DT	04-DEC-2003	(first entry)	
XX	Human PRLR gene.		
XX	human; dg; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;		
KW	cancer; neoplasm; adenocarcinoma; sarcoma; gene.		
XX	Homo sapiens.		
OS	WO2003008583-A2.		
XX	30-JAN-2003.		
XX	26-DEC-2001; 2001WO-US051291.		
XX	02-MAR-2001; 2001US-00798586.		
PR	23-OCT-2001; 2001US-00004113.		
PR	08-NOV-2001; 2001US-00052482.		
PR	30-NOV-2001; 2001US-00997722.		
PR	20-DEC-2001; 2001US-00034650.		
XX	(SAGR-) SAGRES DISCOVERY.		
XX	Morris DW, Engelhard EK;		
PI	WPI; 2003-239337/23.		
XX	New recombinant nucleic acid, useful for treating carcinomas, lymphomas,		
PT	cancers, neoplasm, adenocarcinoma, or sarcomas.		
XX	Claim 1; SEQ ID NO 238; 2304pp; English.		
PS	The invention relates to a novel recombinant nucleic acid comprising a		
XX	nucleotide sequence selected from any of the 660 sequences fully defined		
CC	in the specification. A polynucleotide of the invention has cytostatic		
CC	activity, and may have a use in gene therapy, or in a vaccine. The		
CC	recombinant nucleic acids and polypeptides are useful for treating		
CC	carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and		
CC	sarcomas. The present sequence represents a human gene of the invention.		
XX	Sequence 96599 BP; 28791 A; 19399 C; 19931 G; 28478 T; 0 U; 0 Other;		
SQ	Query Match 27.9%; Score 279.4; DB 10; Length 96599;		
	Best Local Similarity 90.6%; Pred. No. 7.4e-59;		
	Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;		
QY	5	GAAGAAAGCACATCTCGGCGGCGGTGGCTCACACCTGTAATCCCAACACTTTGGGA	64
Db	49044	GAAAAATAACACATCTGGGCGGCGGTGGCTCACGCTGTAATCCAGCACACTTTGGGA	49103
QY	65	GGCGAGGGCGGTGGGATCACTAGTTCAGGAGATGGAGACCATCTCGGTGACACGGTGAA	124
Db	49104	GGCGAGGGCGGCGGATCAGAGTTCAGAGATCGAGACCATCCCGGCTTAAACGGTGAA	49163
QY	125	ACCTGTCTCTACTAAAAACAAAAAATTAGCCGGCGGTGGTGACGCGCTCTAGTC	184
Db	49164	ACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGCGGTAGTGGCGGCGCTGTAGTC	49223
QY	185	CCAGCTTACTCGGAGGCTGAGGCAGAGAGATGGCTGAACCCGGAGCGGAGACTTGAG	244
Db	49224	CCAGCTTACTTGGAGGCTGAGGCAGAGAGATGGCGTGAACCCGGAGCGGAGCTTGCAG	49283
QY	245	TGAGCGGAGGTTGCACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAA	304
Db	49284	TGAGCGGAGATTGCGCACTGCGACTCCAGCTTGGGCGCAGAGAGACTCCGCTCTCAA	49343
QY	305	AAAAAAAAAAGAAAAAGAAATCACA	333
Db	49344	AAAAAAAAAAGAAAAAGAAATCACA	49372
RESULT 10			
ADE95920			

Db	49344	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	49372	
RESULT 11				
ADL15049				
ID	ADL15049	standard; DNA; 77781 BP.		
XX	XX			
AC	ADL15049;			
XX	XX			
DT	06-MAY-2004	(first entry)		
XX	XX			
DE	Human melanoma associated MAGE-like	DNA for cancer treatment.		
XX	XX			
KW	ds; gene; cytostatic; gene therapy; binding moiety; medicine; imaging;			
KW	diagnosis; prognosis; mantle cell lymphoma; cancer.			
XX	XX			
OS	Homo sapiens.			
XX	XX			
PN	WO2003068268-A2.			
XX	XX			
PD	21-AUG-2003.			
XX	XX			
PF	13-FEB-2003; 2003WO-EP001461.			
XX	XX			
PR	14-FEB-2002; 2002GB-00003480.			
XX	XX			
PR	29-JUN-2002; 2002GB-00015095.			
XX	XX			
PA	(BIOI-) BIOINVENT INT AB.			
XX	XX			
PI	Ek S, Borrebaeck CAK, Ehinger M;			
XX	XX			
DR	WPI; 2003-697496/66.			
XX	XX			
DR	P-PSDB; ADL15047.			
XX	XX			
PT	New compound for treating, imaging, diagnosing or prognosing mantle cell			
PT	lymphoma, comprises a binding moiety (e.g. antibody) that binds to a			
PT	protein (e.g. human autotoxin polypeptide), and a further moiety (e.g.			
PT	nucleic acid).			
XX	XX			
PS	Disclosure; SEQ ID NO 61; 342pp; English.			
XX	XX			
CC	The invention relates to a compound comprising a binding moiety which			
CC	selectively binds to a protein or polypeptide listed in the specification			
CC	(e.g. human autotoxin polypeptide or human CD24 signal transducer			
CC	polypeptide), and a further moiety. The compound is useful in medicine or			
CC	in the treatment, imaging, diagnosis or prognosis of mantle cell			
CC	lymphomas (MCL). It is used in preparing a medicament for treating MCL, a			
CC	diagnostic or prognostic agent for MCL, or an agent for imaging MCL cells			
CC	in the body of an individual. This sequence corresponds to a gene			
CC	encoding one of the polypeptides of the invention.			
XX	XX			
SEQ	Sequence 77781 BP; 20855 A; 16833 C; 18350 G; 21743 T; 0 U; 0 Other;			
Query Match	27.9%;	Score 278.8; DB 10; Length 77781;		
Best Local Similarity	89.4%;	Pred. No. 9.7e-59;		
Matches	301; Conservative	0; Mismatches 37; Indels 0; Gaps 0;		
QY	1	AACAGAGAAGACATCTCGCGCGCGCGTGGCTCACACCTGTAAATCCCAACTTTT	60	
Db	60971	ATTCGGAATAGAAACACAGAGCGCGCGCGTGGCTCACACCTGTAAATCCCAACTTTT	61030	
QY	61	GGGAGCGCGAGCGGGTGGATCACTAGGTGAGAGATGGAGACCATCTGGCTGCACACGG	120	
Db	61031	GGGAGCGCGAGCGGGTGGATCACTAGGTGAGAGATGGAGACCATCTGGCTGCACACGG	61090	
QY	121	TGAACCCCTGTCTCTACTATAAAAAACAAAAAATAGCCGGCGTGGTGGCGCGCTGT	180	
Db	61091	TGAACCCCTGTCTCTACTATAAAAAATACAAAAATAGCTGGCGTGGTGGCGCGCTGT	61150	
QY	181	AGTCCAGCTACTCGGAGGCTGAGGAGAGATGGCTGACCCGAGGCGGAACTT	240	
Db	61151	AGTCCAGCTACTCGGAGGCTGAGGAGAGATGGCTGACCCGAGGCGGAACTT	61210	

```
Db 29024 AGTCCATCTCGCGCGCGGTCGCTGCTCAGCTGTATCCAGCACTTTGGGAGGCCGA 29083
Qy 71 GCGCGGTGGATCACTAGTGTCAAGAGATGGAGACCATCTTGCTGACACGGTGAAACCCCTG 130
Db 29084 GCGCGCGCGATCACGAGTCAAGAGATCGAGACCATCTTGCTGCTAACACGGTGAAACCCCG 29143
Qy 131 TCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCTGTAGTCCAGCT 190
Db 29144 TCTCTACTAAAAAATACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCTGTAGTCCAGCT 29203
Qy 191 ACTCGGAGGCTGAGCGCAGGAGAAATGGCTGAAACCCGGGAGCGGAACTTGCAGTGAACC 250
Db 29204 ACTCGGAGGCTGAGCGCAGGAGAAATGGCTGAAACCCGGGAGCGGAGCTTGCAGTGAACC 29263
Qy 251 GAGTTGACCACTGCTCATCTCCAGCTCGGCGCAACACAGTGAGACTCCGCTCTCAAAAAAAA 310
Db 29264 GAGATCGCGCCACTGCTCCAGCTCGGCGCAACAGAGAGACTCCGCTCTCAAAAAAAA 29323
Qy 311 AAAAGAAAGAAAGAAATACATCTCATTCAGTGGTGGCATTTAAACTATT 364
Db 29324 AAAAGAAAGAAAGTAGTCCATCTCTCTGCTCTGAAAGAGATTATAGAAATT 29377

RESULT 13
AAK90223
ID AAK90223 standard; DNA; 3407 BP.
XX
AC AAK90223;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human digestive system antigen genomic sequence SEQ ID NO: 3799.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001324.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
```









PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.

(ROSE/) ROSEN C A.  
(RUBE/) RUBEN S M.  
(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-625420/59.

New nucleic acid molecule, useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g. cancer, liver disorders or neural disorders.

Disclosure; SEQ ID NO 749; 216pp; English.

The invention relates to an isolated nucleic acid molecule encoding a polypeptide. The nucleic acid is useful for preparing a medicament for preventing, treating or ameliorating a medical condition e.g. cancer, liver disorders such as hepatitis or neural disorders such as Alzheimer's disease. The present sequence represents a human novel colon related polypeptide DNA. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at [seqdata.uspto.gov/sequence.html?DocID=20030050231](http://seqdata.uspto.gov/sequence.html?DocID=20030050231).

Sequence 3407 BP; 1134 A; 634 C; 666 G; 973 T; 0 U; 0 Other;

Query Match 27.8%; Score 278.2; DB 9; Length 3407;

Best Local Similarity 91.3%; Pred. No. 4.6e-59;

Matches 295; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 6 AAGAAAGCACATCTCGCGCGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGGAG 65  
Db 3068 AAAAAAGTGAACAACGGCGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGGAG 3127

Qy 66 GCCGAGGCGGTGATCATCTAGGTGAGGATGGAGACATCTTGCTGACACGGTGAAA 125  
Db 3128 GCCGAGGCGGTGATCATCTAGGTGAGGATGGAGACATCTTGCTGACACGGTGAAA 3187

Qy 126 CCCTGTCTCTAAACACAAAAAATAGCCGGCGGTGGCGAGCGGCTGTAGTCC 185  
Db 3188 CCCCGTCTCTAAACACAAAAAATAGCCGGCGGTGGCGAGCGGCTGTAGTCC 3247

Qy 186 CAGTTACTCGGAGGCTGAGGAGGAGAAATGGCCCTGAACCCGGGAGGCGGAACCTTGCACT 245  
Db 3248 CAGTTACTCGGAGGCTGAGGAGGAGAAATGGCCCTGAACCCGGGAGGCGGAGCTTGCACT 3307

Qy 246 GAGCCGAGTTTGACCACTGCACCTCCAGCCTGGGCAACACAGTGGAGCTCCGCTCTCAAAA 305  
Db 3308 GAGCCGAGATCGCCCACTGCACCTCCAGCCTGGGCAACACAGTGGAGCTCCGCTCTCAAAA 3367  
Qy 306 AAAAAAAGAAAAAGAAAAAGAAA 328  
Db 3368 AAAAAAAGAAAAAGAAAAAGAAA 3390

Search completed: July 4, 2005, 03:03:31  
Job time : 358.271 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:41:10 ; Search time 113.609 seconds  
(without alignments)  
14417.061 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_8000\_9000  
Perfect score: 1001  
Sequence: 1 aacagaagaagcagcatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	10818	4	US-09-949-016-13583
c 2	600.6	60.0	601	4	US-09-949-016-63411
c 3	600.6	60.0	601	4	US-09-949-016-63412
c 4	599.4	59.9	601	4	US-09-949-016-63410
c 5	533.6	53.3	601	4	US-09-949-016-63409
c 6	518.6	51.8	601	4	US-09-949-016-63408
c 7	470.6	47.0	601	4	US-09-949-016-63407
c 8	465.6	46.5	601	4	US-09-949-016-63406
c 9	280.2	28.0	84252	4	US-09-949-016-17315
c 10	279.4	27.9	169334	4	US-09-949-016-15999
c 11	278.4	27.8	38052	4	US-09-949-016-13825
c 12	277.6	27.7	130724	4	US-09-949-016-13753
c 13	277.4	27.7	108341	4	US-09-949-016-16621
c 14	277.4	27.7	115814	4	US-09-949-016-16205
c 15	276.8	27.7	601	4	US-09-949-016-130842
c 16	276.8	27.7	20495	4	US-09-949-016-17198
c 17	276.8	27.7	114793	4	US-10-148-806-3
c 18	276.8	27.7	134292	4	US-09-949-016-12158
c 19	276.6	27.6	121982	4	US-09-949-016-12085
c 20	276.6	27.6	121982	4	US-09-949-016-14105
c 21	276.4	27.6	181429	4	US-09-949-016-12372
c 22	276.4	27.6	181430	4	US-09-949-016-15772
c 23	276.2	27.6	84425	4	US-09-949-016-17402
c 24	275.2	27.5	133719	4	US-09-949-016-15092
c 25	275	27.5	68778	4	US-09-949-016-15406
c 26	274.6	27.4	23766	4	US-09-949-016-13569
c 27	274.4	27.4	636591	4	US-09-949-016-11808

c 28	274.4	27.4	636591	4	US-09-949-016-13388	Sequence 13388, A
c 29	274.2	27.4	14543	4	US-09-959-873B-17	Sequence 17, Appl
c 30	273.4	27.3	3243	4	US-09-949-016-12186	Sequence 12186, A
c 31	273.4	27.3	60465	4	US-09-949-016-15995	Sequence 15995, A
c 32	273.4	27.3	76264	4	US-09-949-016-15773	Sequence 15773, A
c 33	273.2	27.3	14636	3	US-09-173-914-6	Sequence 6, Appl
c 34	273	27.3	192956	4	US-09-949-016-14382	Sequence 14382, A
c 35	272.8	27.3	154023	4	US-09-949-016-17057	Sequence 17057, A
c 36	272.6	27.2	15392	4	US-09-949-016-13596	Sequence 13596, A
c 37	272.6	27.2	48316	4	US-09-949-016-12853	Sequence 12853, A
c 38	272.6	27.2	48316	4	US-09-949-016-13405	Sequence 13405, A
c 39	272.6	27.2	104428	4	US-09-949-016-12737	Sequence 12737, A
c 40	272.6	27.2	104429	4	US-09-949-016-13814	Sequence 13814, A
c 41	272.4	27.2	33379	4	US-09-949-016-13861	Sequence 13861, A
c 42	272.4	27.2	118382	4	US-09-949-016-15996	Sequence 15996, A
c 43	272.4	27.2	118382	4	US-09-949-016-15997	Sequence 15997, A
c 44	272.2	27.2	601	4	US-09-949-016-22105	Sequence 22105, A
c 45	272.2	27.2	601	4	US-09-949-016-154105	Sequence 154105,

ALIGNMENTS

RESULT 1

US-09-949-016-13583  
; Sequence 13583, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13583  
; LENGTH: 10818  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13583

Query Match	100.0%	Score	1001;	DB	4;	Length	10818;
Best Local Similarity	100.0%;	Pred. No.	6.5e-257;	Mismatches	0;	Indels	0;
Matches	1001;	Conservative	0;				
Qy	1	AACAGAGAAAGACATCTCGCGCGGGGTGGTGGCTCACACCTGTATCCACACATTT	60				
Db	5245	AACAGAGAAAGACATCTCGCGCGGGGTGGTGGCTCACACCTGTATCCACACATTT	5304				
Qy	61	GGGAGCGCGAGCGGGTGGATCACTAGGTACAGAGATGAGACCATCTCGTGCACCG	120				
Db	5305	GGGAGCGCGAGCGGGTGGATCACTAGGTACAGAGATGAGACCATCTCGTGCACCG	5364				
Qy	121	TGAACCCCTGTCTCTACTATAAAACACAAAAATAGCCGGCGGTGGTGGCAGGCCCTGT	180				
Db	5365	TGAACCCCTGTCTCTACTATAAAACACAAAAATAGCCGGCGGTGGTGGCAGGCCCTGT	5424				
Qy	181	AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTCAACCCGGGAGCGGAATTT	240				
Db	5425	AGTCCAGCTACTCGGAGGCTGAGCGAGGAGATGGCTCAACCCGGGAGCGGAATTT	5484				
Qy	241	GCAGTGAGCCGAGGTTCACACCTGCATCCAGCCTGGGCAACACAGTGAGACTCCGTCT	300				
Db	5485	GCAGTGAGCCGAGGTTCACACCTGCATCCAGCCTGGGCAACACAGTGAGACTCCGTCT	5544				
Qy	301	CAAAAAAAGAAAAAGAAAAAATCATCTCATTCAGTGGTGGCATTTAAAAAC	360				

Db	5545	CAAAAAAAAAAAGAAAGAAAGAAATCAATCTATTCAAGTGGTGGCAATTTAAAAAC	5604
Qy	361	TATTTAGCCTTCTGTAGGCAAGGTTAGTATCTTGTTTTTCCAGACCTCAAGGTGTTTTT	420
Db	5605	TATTTAGCCTTCTGTAGGCAAGGTTAGTATCTTGTTTTTCCAGACCTCAAGGTGTTTTT	5664
Qy	421	TTGTTTGTTTTTTCATACCGGTGTGTGTCGGGTGTGGCCACTAAAAGCTCAAGCAAG	480
Db	5665	TTGTTTGTTTTTTCATACCGGTGTGTGTCGGGTGTGGCCACTAAAAGCTCAAGCAAG	5724
Qy	481	AAATPAAACAACTACAACAATCTAAATACCAATAGTATAAAAATAATAGCATCTGGCTA	540
Db	5725	AAATPAAACAACTACAACAATCTAAATACCAATAGTATAAAAATAATAGCATCTGGCTA	5784
Qy	541	ATTGCTGGACACTGTTTTAAAGTGGTTGTCATGCCCTCAGCTCAATTAACCTATTACCTGTT	600
Db	5785	ATTGCTGGACACTGTTTTAAAGTGGTTGTCATGCCCTCAGCTCAATTAACCTATTACCTGTT	5844
Qy	601	ATTATTTGGCCCTATTTTTCAAAACAAGAGCCCAAGGCTCAGACAGTTAACTAACAGCCTC	660
Db	5845	ATTATTTGGCCCTATTTTTCAAAACAAGAGCCCAAGGCTCAGACAGTTAACTAACAGCCTC	5904
Qy	661	TCAAAAGAAACTCTGCAGAGATATTTAAATTTTAAATAATCAGAGAAATTTAAACACAA	720
Db	5905	TCAAAAGAAACTCTGCAGAGATATTTAAATTTTAAATAATCAGAGAAATTTAAACACAA	5964
Qy	721	GAAAGTTGAAATTTAGAGGTCACAGCAGCTAAGCTTGTTTGCTTTTGAAACAGTGTCTGCT	780
Db	5965	GAAAGTTGAAATTTAGAGGTCACAGCAGCTAAGCTTGTTTGCTTTTGAAACAGTGTCTGCT	6024
Qy	781	ACTGGGAAAAAGGCAAGTCTTTGGCTTTTCTTAATAATATGATACAAGACTCTGTAAATCAT	840
Db	6025	ACTGGGAAAAAGGCAAGTCTTTGGCTTTTCTTAATAATATGATACAAGACTCTGTAAATCAT	6084
Qy	841	ATTTTTCATGCTGTAAGCTAGAAATGAGCCGGTGCATGGCAGATGCCAGTAATCCC	900
Db	6085	ATTTTTCATGCTGTAAGCTAGAAATGAGCCGGTGCATGGCAGATGCCAGTAATCCC	6144
Qy	901	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGGAGCTCAGAGTTCAGAGCCAGCCTG	960
Db	6145	AGCACTCTGGGAGACTGAAGTGGGAAGATCACTTGGAGCTCAGAGTTCAGAGCCAGCCTG	6204
Qy	961	GGCAACTAAAAATTTAAAAATAAAAAATCAATTTGTTTTT	1001
Db	6205	GGCAACTAAAAATTTAAAAATAAAAAATCAATTTGTTTTT	6245

RESULT 2  
 US-09-949-016-63411/c  
 ; Sequence 63411, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 63411  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-63411

Query Match	60.8%;	Score 600.6;	DB 4;	Length 601;
Best Local Similarity	99.8%;	Pred. No. 1.3e-150;		
Matches 600;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	29	GTGGTGCTC	CACACCTGTAATCCCAACACATTTGGGAGCGGAGCGGGTGGATC	ACTAGG 88
DB	601	GTGGTGCTC	ACACCTGTATCCCAACATTTGGGAGCGGAGCGGGTGGATC	ACTAGG 542
QY	89	TCAGGAGAT	GGAGACCAATCCTGGCTGACACGGTGAAACCCCTGTCTCTATAAAACACAA	148
DB	541	TCAGGAGAT	GGAGACCAATCCTGGCTGACACGGTGAAACCCCTGTCTCTATAAAACACAA	482
QY	149	AAATTAAT	TAGCGGCGTGGTGCAGGGCCCTGTAGTCCAGCTACTCTCGGAGGCTGAGGCA	208
DB	481	AAATTAAT	TAGCGGCGTGGTGCAGGGCCCTGTAGTCCAGCTACTCTCGGAGGCTGAGGCA	422
QY	209	GGAGAAT	TGGCTGAACCCGGGAGCGGAACTTGCAGTGAGCCGAGGTTGCACCACTGCAC	268
DB	421	GGAGAAT	TGGCTGAACCCGGGAGCGGAACTTGCAGTGAGCCGAGGTTGCACCACTGCAC	362
QY	269	TCCAGCCT	GGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAA	328
DB	361	TCCAGCCT	GGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAGAAAGAAAA	302
QY	329	TCAATCT	CATTCAAGTGGTGGCAATTTAAACATATTTAGCTCTTCTCTAGGCAAGGTTAG	388
DB	301	KCAATCT	CATTCAAGTGGTGGCAATTTAAACATATTTAGCTCTTCTCTAGGCAAGGTTAG	242
QY	389	TAICTTGT	TTTTTCAGACCTCAAGGTGTTTTTTGTTTTTTTTCATACCGGTGTGTGG	448
DB	241	TATCTTGT	TTTTTCAGACCTCAAGGTGTTTTTTGTTTTTTTTCATACCGGTGTGTGG	182
QY	449	TCCTGG	TGTGCGCACTAAAAGCTACAGCAAGAAATAATAACACTACAACAATACTAAT	508
DB	181	TCCTGG	TGTGCGCACTAAAAGCTACAGCAAGAAATAATAACACTACAACAATACTAAT	122
QY	509	ACCAATAG	TATAAAAAATAATAGCATCTGGCTAATTGTCTGGACACTGTTTTTAAGTGTTG	568
DB	121	ACCAATAG	TATAAAAAATAATAGCATCTGGCTAATTGTCTGGACACTGTTTTTAAGTGTTG	62
QY	569	CATGCC	CTCAGCTCAATTAACCTCATTTACTCTGTATTATTGGCCCTCAATTTTACAACRAGGA	628
DB	61	CATGCC	CTCAGCTCAATTAACCTCATTTACTCTGTATTATTGGCCCTCAATTTTACAACRAGGA	2
QY	629	G	629	
DB	1	G	1	

```

RESULT 3
US-09-949-016-63412/c
; Sequence 63412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63412
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63412

```



		Query Match	53.3%	Score 533.6	DB 4	Length 601		
		Best Local Similarity	99.8%	Pred. No. 9.6e-133				
		Matches 533	Conservative 1	Mismatches 0	Indels 0	Gaps 0		
QY	1	AACAGAGAAGACATCTCGCGCGGGTGGTGCACACACCTGTATCCCAACACTTT	60					
Db	534	AACAGAAGAAGCACATCTCGCGCGGGGTGGTGCACACCTGTATCCCAACACTTT	475					
QY	61	GGGAGGCCGAGCGGGTGGATCACTAGGTCAAGAGATGGAGACCATCTCGGTGCACACGG	120					
Db	474	GGGAGGCCGAGCGGGTGGATCACTAGGTCAAGAGATGGAGACCATCTCGGTGCACACGG	415					
QY	121	TGAACCCCTGTCTCTACTAAACACACAAAAATTAGCCGGGGTGGTGGCAGGCCCTGT	180					
Db	414	TGAACCCCTGTCTCTACTAAACACACAAAAATTAGCCGGGGTGGTGGCAGGCCCTGT	355					
QY	181	AGTCCCAGCTACTCTCGGAGGCTGAGGAGGAGAATGGCTCAACCCGGCGAGCGGAACTT	240					
Db	354	AGTCCCAGCTACTCTCGGAGGCTGAGGAGGAGAATGGCTCAACCCGGCGAGCGGAACTT	295					
QY	241	GCAGTAGCCGAGGTTGCAACACTTGCACTCCAGCCTCGGGCAACACAGTGAGACTCCGCT	300					
Db	294	GCAGTAGCCGAGGTTGCAACACTTGCACTCCAGCCTCGGGCAACACAGTGAGACTCCGCT	235					
QY	301	CAAAAAAAGAAAAAGAAAAAGAAATCACATCTCAATCAAGTGGTGGCATTTAAAC	360					
Db	234	CAAAAAAAGAAAAAGAAAAAGAAATCACATCTCAATCAAGTGGTGGCATTTAAAC	175					
QY	361	TATTTAGCCTTCTGTAGGCAAGGTTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT	420					
Db	174	TATTTAGCCTTCTGTGTAGGCAAGGTTAGTATCTTGTGTTTTTCCAGACCTCAAGGTGTTTT	115					
QY	421	TTGTTTGTGTTTTTTCATACCGGTGTGTGCTGTGGGTGCGGCACATAAAGCTACAAGCAAG	480					
Db	114	TTGTTTGTGTTTTTTCATACCGGTGTGTGCTGTGGGTGCGGCACATAAAGCTACAAGCAAG	55					
QY	481	AAATAATAACAACCTACAACAATACTAATACCAATAGTATATAAAAAATAATAGCATC	534					
Db	54	AAATAATAACAACCTACAACAATACTAATACCAATAGTATATAAAAAATAATAGCATC	1					

```

RESULT 6
US-09-949-016-63408/c
; Sequence 63408, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63408
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63408

```

Qy	1	AA	CAGAAGAAG	CACATCTCGGCGGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT	60
Db	519	AA	CAGAAGAAG	CACATCTCGGCGGGCGTGGTGGCTCACACCTGTAAATCCCAACACTTT	460
Qy	61	GG	GAGGCCGAGGCGGGTGGATCACTAGCTCAGAGATGGAGACCATCTCTGGCTGACACGG	120	
Db	459	GG	GAGGCCGAGGCGGGTGGATCACTAGCTCAGAGATGGAGACCATCTCTGGCTGACACGG	400	
Qy	121	TG	AAACCCCTGTCTCTACTAAACACACAAAAAATTAGCCGGGCGTGGTGGCAGGGCGCTGT	180	
Db	399	TG	AAACCCCTGTCTCTACTAAACACACAAAAAATTAGCCGGGCGTGGTGGCAGGGCGCTGT	340	
Qy	181	AGTCC	CAGCTACTCGGGAGGCTGAGGCAGAGAAATGCGTGAACCCGGGAGGCGGAACCT	240	
Db	339	AGTCC	CAGCTACTCGGGAGGCTGAGGCAGAGAAATGCGTGAACCCGGGAGGCGGAACCT	280	
Qy	241	GCAGT	GAGCGAGGTTGCAACCATGTCACTCCAGCTCGGGCAACACAGTGAAGACTCCGTCT	300	
Db	279	GCAGT	GAGCGAGGTTGCAACCATGTCACTCCAGCTCGGGCAACACAGTGAAGACTCCGTCT	220	
Qy	301	CA	AAAAAAGAAAAAGAAAAAGAAATCACATCTCATTCAAGTGGTGGCATTTAAAC	360	
Db	219	CA	AAAAAAGAAAAAGAAAAAGAAATCACATCTCATTCAAGTGGTGGCATTTAAAC	160	
Qy	361	TATTTAG	CCCTTCTGTAGGCAAGTTAGTATCTTTGTTTTTCCAGACCTCAAGGTGTTTT	420	
Db	159	TATTTAG	CCCTTCTGTAGGCAAGTTAGTATCTTTGTTTTTCCAGACCTCAAGGTGTTTT	100	
Qy	421	TTGTTT	TTTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAGCTTACAGCAAG	480	
Db	99	TTGTTT	TTTTTTTCATACCGGTGTGGTCTGGGTGGCCACTAAAGCTTACAGCAAG	40	
Qy	481	AAATAA	TAAACAACCTACAACTAAATAACAATAGTAT	519	
Db	39	AAATAA	TAAACAACCTACAACTAAATAACAATAGTAT	1	

```

RESULT 7
US-09-949-016-63407/c
; Sequence 63407, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 63407
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63407

```

	Query Match	47.08;	Score	470.6;	DB	4;	Length	601;
	Best Local Similarity	99.8;	Pred. No.	6.3e-116;				
	Matches	470;	Conservative	1;	Mismatches	0;	Indels	0;
								Gaps
								0;
Qy	1	AACAGAAAGAACACATCTCGGCGGGCGTGGTGGCTCACACCTGTGTAATCCCAACACTTT	60					
Db	471	AACAGAAAGAACACATCTCGGCGGGCGTGGTGGCTCACACCTGTGTAATCCCAACACTTT	412					
Qy	61	GGGAGGCCGAGGCGGGTGGATCACTAGTCCAGGAGATGGAGACCATCTCTGGCTGACACGG	120					
Db	411	GGGAGGCCGAGGCGGGTGGATCACTAGTCCAGGAGATGGAGACCATCTCTGGCTGACACGG	352					



Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCGGCGGTGGTGGAGCGCCTGT 180  
Db |||||  
Qy 351 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCGGCGGTGGTGGAGCGCCTGT 292  
Db |||||  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGGCTGAACCCGGAGCGCGAACTT 240  
Db |||||  
Qy 291 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGGCTGAACCCGGAGCGCGAACTT 232  
Db |||||  
Qy 241 GCAGTGAGCCGAGGTGACACCTGCACCTCAGCTCGGCGAGGAGAAACAGTGAAGCTCCGCT 300  
Db |||||  
Qy 231 GCAGTGAGCCGAGGTGACACCTGCACCTCAGCTCGGCGAGGAGAAACAGTGAAGCTCCGCT 172  
Db |||||  
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCACTTAAAC 360  
Db |||||  
Qy 171 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCACTTAAAC 112  
Db |||||  
Qy 361 TATTTAGCCTTTCTGTAGGCAAGGTAGTATCTTTTCCAGACCTCAAGTGTGTTTT 420  
Db |||||  
Qy 111 TATTTAGCCTTTCTGTAGGCAAGGTAGTATCTTTTCCAGACCTCAAGTGTGTTTT 52  
Db |||||  
Qy 421 TTGTTGTTTTTTCATACCGGTGTGTGCTGTGGGTGGCGCACTAAAGCT 471  
Db |||||  
Qy 51 TTGTTGTTTTTTCATACCGGTGTGTGCTGTGGGTGGCGCACTAAAGCT 1  
Db |||||  
RESULT 8  
US-09-949-016-63406/c  
; Sequence 63406, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63406  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-63406  
Query Match 46.5%; Score 465.6; DB 4; Length 601;  
Best Local Similarity 99.8%; Pred. No. 1.4e-114;  
Matches 465; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AACAGAGAAAGCACTCTCGGCGGCGGTGGTGGCTCACACCTGTATATCCCAACACTTT 60  
Db |||||  
Qy 466 AACAGAGAAAGCACTCTCGGCGGCGGTGGTGGCTCACACCTGTATATCCCAACACTTT 407  
Db |||||  
Qy 61 GGGAGGCGGAGCGGTGGATCACTAGTGTAGGAGATGGAGACCATCTGCTGCACACGG 120  
Db |||||  
Qy 406 GGGAGGCGGAGCGGTGGATCACTAGTGTAGGAGATGGAGACCATCTGCTGCACACGG 347  
Db |||||  
Qy 121 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCGGCGGTGGTGGAGCGCCTGT 180  
Db |||||  
Qy 346 TGAACCCCTGCTCTACTAAAAACACAAAAAATAGCGGCGGTGGTGGAGCGCCTGT 287  
Db |||||  
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGGCTGAACCCGGAGCGCGAACTT 240  
Db |||||  
Qy 286 AGTCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGGCTGAACCCGGAGCGCGAACTT 227  
Db |||||  
Qy 241 GCAGTGAGCCGAGGTGACACCTGCACCTCAGCTCGGCGAGGAGAAACAGTGAAGCTCCGCT 300  
Db |||||

Db 226 GCAGTGAGCCGAGGTGACCACTGCACCTCAGCGCTGGGCAACACAGTGAAGCTCCGCTCT 167  
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCACTTAAAC 360  
Db |||||  
Qy 166 CAAAAAAGAAAAAGAAAAAGAAAAATACATCTCAATCAAGTGGTGGCACTTAAAC 107  
Db |||||  
Qy 361 TATTTAGCCTTTCTGTAGGCAAGGTAGTATCTTTTCCAGACCTCAAGTGTGTTTT 420  
Db |||||  
Qy 106 TATTTAGCCTTTCTGTAGGCAAGGTAGTATCTTTTCCAGACCTCAAGTGTGTTTT 47  
Db |||||  
Qy 421 TTGTTGTTTTTTCATACCGGTGTGTGCTGTGGGTGGCGCACTAA 466  
Db |||||  
Qy 46 TTGTTGTTTTTTCATACCGGTGTGTGCTGTGGGTGGCGCACTAA 1  
Db |||||  
RESULT 9  
US-09-949-016-17315/c  
; Sequence 17315, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17315  
; LENGTH: 84252  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17315  
Query Match 28.0%; Score 280.2; DB 4; Length 84252;  
Best Local Similarity 91.4%; Pred. No. 4e-64;  
Matches 297; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
Qy 4 AGAAGAAAGCACTCTCGGCGGCGGTGGTGGCTCACACCTGTATATCCCAACACTTTGGG 63  
Db |||||  
Qy 35402 ATAAGAGAGAAATACAGGCGGCGGTGGCTCACGCTGTATATCCAGCACTTTGGG 35343  
Db |||||  
Qy 64 AGGCGAGGCGGTGGATCACTAGTGTAGGAGATGGAGACCATCTGCTGCACACGGTGA 123  
Db |||||  
Qy 35342 AGGCGAGGCGGTGGATCACTAGTGTAGGAGATGGAGACCATCTGCTGCACACGGTGA 35283  
Db |||||  
Qy 124 AACCTCTCTCTACTAAAAACACAAAAAATAGCGGCGGTGGTGGAGCGCCTGTAGT 183  
Db |||||  
Qy 35282 AACCTCTCTCTACTAAAAATACAAAAATAGCCGCGGTGGTGGAGCGCCTGTAGT 35223  
Db |||||  
Qy 184 CCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGCCTGAACCCGGAGCGCGAACTTGA 243  
Db |||||  
Qy 35222 CCCAGCTACTCGGAGGCTGAGGCGAGGAGAAATAGCCTGAACCCGGAGCGCGAGCTGA 35163  
Db |||||  
Qy 244 GTGAGCGGAGGTGACCACTGCACCTCAGCTCGGCGAGGAGAAACAGTGAAGCTCCGCTCAA 303  
Db |||||  
Qy 35162 GTGAGCGGAGATCGGCACTGCACCTCAGCTCGGCGAGGAGAAACAGTGAAGCTCCGCTCAA 35103  
Db |||||  
Qy 304 AAAAAAAGAAAAAGAAAAAGAA 328  
Db |||||  
Qy 35102 AAAAAAAGAAAAAGAAAAAGAA 35078  
Db |||||  
RESULT 10  
US-09-949-016-15999  
; Sequence 15999, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15999  
; LENGTH: 169334  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(169334)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15999

Query Match 27.9%; Score 279.4; DB 4; Length 169334;  
Best Local Similarity 90.6%; Pred. No. 9.1e-64;  
Matches 298; Conservative 0; Mismatches 31; Indels 0; Gaps 0;  
  
Qy 5 GAAGAAAGACATCTCGCGCGGCGTGGTCTCACACCTGTATATCCCAACACTTTGGGA 64  
Db 41044 GAAATAACATCTGGCGCGGCGGTGCTCAAGCTGTATATCCCAACACTTTGGGA 41103  
  
Qy 65 GGCGAGCGGTGGATGACTAGTCTAGGAGATGGAGACCATCTGGCTGACACGGTGAA 124  
Db 41104 GGCGAGCGCGGCGATCACAGTCTAGGAGATGGAGACCATCTGGCTGACACGGTGAA 41163  
  
Qy 125 ACCCTGTCTCTACTTAAACACAAAAATAGCGCGGTGGTGGAGACCATCTCTGTAGTC 184  
Db 41164 ACCCGTCTCTACTTAAATACAAAAATAGCGCGGTGGTGGAGACCATCTCTGTAGTC 41223  
  
Qy 185 CCAGTACTCGGAGGTGAGCGAGGAATGGCTGAAACCCGGAGCGGAACTTGGAG 244  
Db 41224 CCAGTACTTGGAGGTGAGCGAGGAATGGCTGAAACCCGGAGCGGAGCTTGGAG 41283  
  
Qy 245 TGAGCGGAGTTGACCACTCCAGCTCCAGCTGGCGACACAGTCCGACTCCGCTCAAA 304  
Db 41284 TGAGCGGAGTTGCGCCACTCCAGCTCCAGCTGGCGACACAGTCCGACTCCGCTCAAA 41343  
  
Qy 305 AAAAAAAGAAAGAAAGAAATTAAGCGGCGGTGGTGGAGACCATCTCTGTAGTC 333  
Db 41344 AAAAAAAGAAAGAAAGAAATTAAGCGGCGGTGGTGGAGACCATCTCTGTAGTC 41372

RESULT 11  
US-09-949-016-13825  
; Sequence 13825, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13825  
; LENGTH: 38052  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(38052)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13825

Query Match 27.8%; Score 278.4; DB 4; Length 38052;  
Best Local Similarity 58.5%; Pred. No. 8.5e-64;  
Matches 566; Conservative 0; Mismatches 386; Indels 16; Gaps 4;  
  
Qy 5 GAAGAAAGACATCTCGCGCGGCGTGGTGGTCTCACACCTGTATATCCCAACACTTTGGGA 64  
Db 3958 GAGTAAAAAATTTCAAGCGCGGCGGTGGTGGTCTCACACCTGTATATCCCAACACTTTGGGA 4017  
  
Qy 65 GGCGAGCGGTGGATCACTAGGTCTAGGAGATGGAGACCATCTCTGTCTGACACGGTGAA 124  
Db 4018 GGCGAGCGGCGAGATCACAGGTCTAGGAGATGGAGACCATCTCTGTCTGACACGGTGAA 4077  
  
Qy 125 ACCCTGTCTCTACTTAAAAACAAAAATTAGCGGCGGTGGTGGAGACCATCTCTGTAGTC 184  
Db 4078 ACCCGTCTCTACTTAAAAATACAAAAAGTTAGCTGGCGGTGGTGGAGACCATCTCTGTAGTC 4137  
  
Qy 185 CCAGTACTCGGAGGTGAGGAGGAGGAGATGGCTGAAACCCGGAGCGGAACTTGCAG 244  
Db 4138 CCAGTACTCGAGAGTTGAGGAGGAGGAGATGGCTGAAACCCGGAGCGGAGTTGCAG 4197  
  
Qy 245 TGAGCGAGGTTCACACCTGCACTCCAGCTCGGCGCAACACAGTGAGACTCCGCTCAAA 304  
Db 4198 TGAGCGAAGATCGCGCACTGCACTCCAGCTCGGCGCAACACAGTGAGACTCCGCTCAAA 4257  
  
Qy 305 AAAAAAAGAAAGAAAGAAATTAAGTGGTGGAGTCTTAAACACTT 364  
Db 4258 AAAAAAAGAAAGAAATTTCAAAACACA--AAATCAATTAATCTATATAGTGAATCAATTA 4315  
  
Qy 365 TAGCTTTCTGTAGGCAAGGTAGTATCTTGTGTTTTCC-----AGACCTCAAGGTGTTT 418  
Db 4316 CAGGTTTTTGAATACATCTTTACTGTACTTGTACTACTAATATTTAATTTAATTAATA 4375  
  
Qy 419 TTTTGTGTTTTTTCATACCGGTGTGTGTCTGGGTGTGGCCACTAAAAGCTCAAGCA 478  
Db 4376 ATTACATTAATTAATAAATTAATTAATTTAAAAAATCTGCTTATATACCGATGTC 4435  
  
Qy 479 AGAATAATTAACACTACACATACTAATACCAATAGTATATAAATAATAGCATCTGGC 538  
Db 4436 ATTACTCTTTTCTAAGTTCAATATAGTAACAGTAACATGCGGAAACCCCGCTCTTAC 4495  
  
Qy 539 TAATTGCTGGACACTGTTTTAAGTGGTTTGCATCGCTCAGCTCATTAATCACTTACCTG 598  
Db 4496 TAAAGTACAAAAATAGCTGGCGGTGGTAGCGGTGCTGTAAACCTAGCTACTTGGGAG 4555  
  
Qy 599 TTATTATTGGCCCTATTTTTACAAACAGGAGCCAGGCTCAGAGCAGTTAACTAACAGCC 658  
Db 4556 GCAGAAACAGGAATCGCTTGAACCTGGAGCGCAGAGGTTGCACTGAGCCGAGATCGGC 4615  
  
Qy 659 TCTCAAAAGAAACCTCGCAGAGATATTAATTTAAAAAATAATGAGAAATTAACAC 718  
Db 4616 CAGTGCAGGACTCCGCTTTTCAAAAAAAGAAAAAAGAAAAAAGAAAAAGTAAAGAAA 4675  
  
Qy 719 AAGAAAGTTGAAATTTAGAGGTACAGGACGCTAAGCTTGTGTTGTTTGAACAGTGTCTG 778  
Db 4676 AATATTTTATGAA-----AGATGCTCTGTGTTAATTTGTTATGTAACAATAATTTTTC 4730  
  
Qy 779 CTACTGGGAAAAAGCGCAAGTCTTGGCTTTTCTAATAATTTGATACAGGACTCTCTGTAATC 838  
Db 4731 CAATTAAGTGGAGGTGCATAAGTATTTCC---AACATTGATCTAAATCCAGTGTAAACA 4787  
  
Qy 839 ATATTTTGCATGCATGTAAGTAAGAAATGAAGCGGCTGCAATGGCATCCAGTATCC 898  
Db 4788 AAAAGTTGATCTTTTATTTAAAACTAAATTTCAAGCGCGGCGCTCTGGCTCACGCTGCAATC 4847



Query Match	27.7%	Score	277.4	DB	4	Length	115814
Best Local Similarity	90.5%	Pred. No.	2.6e-63				
Matches	296	Conservative	0	Mismatches	31	Indels	0
QY	3	CAGAAGAAACACACTCTCGGCGGGCGTGGTGGCTACACCTGTAATCCCAACACTTTGG	62				
Db	8625	CAAAAAAAAAAAGATTTCGGCGCGCGGGTGGCTACGGCTGTAATCCAGACACTTTGG	8566				
QY	63	GAGGCCGAGCGGGTGGATCACTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTG	122				
Db	8565	GAGGCCGAGCGGGCGGATCACGAGGTCAGGAGATCGAGACCATCTCTGGCTAACACGGTG	8506				
QY	123	AAACCTGTCTCTACTTAAAAACAACAAAAATTAGCCGGGCGTGGTGCGAGGGCGCTGTAG	182				
Db	8505	AAACCCCGTCTCTACTTAAAAATAACAAAAATTAGCCGGGCGTGGTAGCGGGCGCTGTAG	8446				
QY	183	TCCACAGTACTCCGGAGGCTGAGCGAGGAAATGGCCCTGAAACCGGAGCGGGAACCTTC	242				
Db	8445	TCCACAGTACTCCGGAGGCTGAGCGAGGAAATGGCCCTGAAACCGGAGCGGAGCTTC	8386				
QY	243	AGTGAGCCGAGGTTGCACCACTGCACCTCCAGCCCTGGGCAACACAGTGAGACTTCGCTCA	302				
Db	8385	AGTGAGCCGAGATCGCGCCACTGCACCTCCAGCCCTGGGCGGACAGAGCCGACTTCGCTCA	8326				
QY	303	AAAAAAAAAAAAAGAAAAAGAAAT	329				
Db	8325	AAAAAAAAAAAAAGAAAAAGAAAT	8299				

RESULT 15  
US-09-949-016-190842/c  
; Sequence 190842, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 190842  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-190842

	Query Match	27.7%	Score 276.8	DB 4	Length 601
	Best Local Similarity	89.0%	Prod. No. 3.3e-64		
	Matches 299	Conservative 0	Mismatches 37	Indels 0	Gaps 0
Qy	1	AACAGAAGAACATCTCGCGCGGCGTGGTGGCTCACACCTGTGTATCCCAACACTTT	60		
Db	471	AATTGAAGATATACAACTTGGCGCGGCGTGGCTCACGCTGTGTATCCGACACTTT	412		
Qy	61	GGGAGCGCGGCGGTGGATCACTAGTCTAGGAGATCGGAGACCATCTCTGCTCGACACGG	120		
Db	411	GGGAGCGCGGCGGCGGATCAAGAGTCAAGAGATCGAGACCATCTCTGCTTAACACGG	352		
Qy	121	TGAAACCTGTCTCTACTTAAAAAAACAAAAAATTAGCCGGCGGTGGTAGGCGCCTGT	180		
Db	351	TGAAACCCCGTCTCTACTTAAAAAAACAAAAAATTAGCCGGCGGTGGTAGGCGCCTGT	292		
Qy	181	AGTCCCAGCTACTCGGAGGCTTGAGGCGAGGAGATGGCTTGAACCCGGGAGCGGAACTT	240		

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 17:45:20 ; Search time 400.727 Seconds  
(without alignments)  
15667.668 Million cell updates/sec

Title: US-09-936-271c-13\_copy\_8000\_9000  
Perfect score: 1001  
Sequence: 1 aacagagaagacacatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283.4	28.3	87467	19	US-10-741-601-5634 Sequence 5634, Ap
2	283.4	28.3	87467	21	US-10-741-600-17624 Sequence 17624, A
3	279.4	27.9	96599	14	US-10-105-948-4 Sequence 4, Appli
4	279.4	27.9	96599	18	US-10-052-482-178 Sequence 178, App
5	278.2	27.8	3407	10	US-09-764-872-749 Sequence 749, App
6	276.8	27.7	114793	15	US-10-148-806-3 Sequence 3, Appli
7	276.8	27.7	114793	22	US-10-859-792-3 Sequence 3, Appli

c 8	276.6	27.6	119501	17	US-10-174-319-15 Sequence 15, Appli
c 9	276.4	27.6	31474	10	US-09-764-891-8149 Sequence 8149, Ap
c 10	276.4	27.6	204621	13	US-10-087-192-958 Sequence 958, App
c 11	276.4	27.6	290367	20	US-10-719-993-6887 Sequence 6887, Ap
c 12	276.2	27.6	100445	19	US-10-322-281-170 Sequence 170, App
c 13	276	27.6	344805	20	US-10-779-271-1 Sequence 1, Appli
c 14	276	27.6	354592	22	US-10-737-082-70 Sequence 70, Appli
c 15	276	27.6	354592	22	US-10-765-790-70 Sequence 70, Appli
c 16	275.6	27.5	86361	19	US-10-741-601-5702 Sequence 5702, Ap
c 17	275.6	27.5	86361	21	US-10-741-600-17803 Sequence 17803, A
c 18	275.4	27.5	109730	21	US-10-741-600-17809 Sequence 17809, A
c 19	275.2	27.5	41104	9	US-09-816-685-3 Sequence 3, Appli
c 20	275.2	27.5	41104	18	US-10-639-708-3 Sequence 3, Appli
c 21	274.8	27.5	2328	10	US-09-764-891-6578 Sequence 6578, Ap
c 22	274.6	27.4	1400	9	US-09-263-959-295 Sequence 295, App
c 23	274.6	27.4	2325	10	US-09-764-891-6579 Sequence 6579, Ap
c 24	274.6	27.4	177380	21	US-10-484-577-683 Sequence 683, App
c 25	274.6	27.4	267156	11	US-09-968-007A-427 Sequence 427, App
c 26	274.6	27.4	267156	21	US-10-843-641A-6897 Sequence 6897, Ap
c 27	274.6	27.4	684973	9	US-09-263-959-1 Sequence 1, Appli
c 28	274.4	27.4	466	20	US-10-357-930-35244 Sequence 35244, A
c 29	274.2	27.4	14543	19	US-10-344-908-1 Sequence 1, Appli
c 30	274.2	27.4	14543	21	US-10-901-583-17 Sequence 17, Appli
c 31	273.6	27.3	392	20	US-10-357-930-40676 Sequence 40676, A
c 32	273.6	27.3	392	20	US-10-357-930-44084 Sequence 44084, A
c 33	273.6	27.3	466	13	US-10-027-632-114651 Sequence 114651, Sequence 114651,
c 34	273.6	27.3	466	17	US-10-027-632-114651 Sequence 114651,
c 35	273.6	27.3	43411	19	US-10-450-826-76 Sequence 76, Appli
c 36	273.4	27.3	67883	22	US-10-737-082-66 Sequence 66, Appli
c 37	273.4	27.3	67883	22	US-10-765-790-66 Sequence 66, Appli
c 38	273.4	27.3	202814	20	US-10-719-993-6812 Sequence 6812, Ap
c 39	273.2	27.3	93443	20	US-10-723-860-2089 Sequence 2089, Ap
c 40	273.2	27.3	276820	18	US-10-271-416-9 Sequence 9, Appli
c 41	273.2	27.3	338702	13	US-10-087-192-292 Sequence 292, App
c 42	273	27.3	303172	20	US-10-719-993-6890 Sequence 6890, Ap
c 43	273	27.3	322885	21	US-10-679-366-3 Sequence 3, Appli
c 44	272.8	27.3	32203	9	US-09-764-869-1849 Sequence 1849, Ap
c 45	272.8	27.3	32203	14	US-10-091-504-1849 Sequence 1849, Ap

ALIGNMENTS

RESULT 1

US-10-741-601-5634  
; Sequence 5634, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5634  
; LENGTH: 87467  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-5634

Query Match 28.3%; Score 283.4; DB 19; Length 87467;  
Best Local Similarity 90.7%; Pred. No. 9.6e-61;  
Matches 302; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1 AACAGAGAAAGACATCTCGCGCGGCGTGTGCTCACCTCTAATCCAACTTT 60  
Db 15709 AAAAGTTGAAGAAAAGATCGCGCGCGGCGTGTGCTTAATCCAGCACTTT 15768

QY 61 GGGAGCCGAGCGGCTGATCAGTCTAGGATCGAGACCATCTGCTGACACGG 120  
Db 15769 GGGAGCCGAGCGGCTGATCAGTCTAGGATCGAGACCATCTGCTGACACGG 15828



	Matches	298;	Conservative	0;	Mismatches	31;	Indels	0;	Gaps	0;
Qy	5	GAAGAAAGCACATCTCGGCGGGCGTGGTGGTTCACACCTGTAATCCCAACACTTTGGGA	64							
Db	49044	GAATAAACACATCTGGGCGGGCGGTGGCTCAGCGCTGTAATCCACGACATTTGGGA	49103							
Qy	65	GGCGGAGGGGGTGGATCACTAGGTGAGAGATGGAGACCATCTTGGCTGACACGGTGAA	124							
Db	49104	GGCGGAGGGGGCGGATCACGAGTTCAGGAGATCGAGACCATCCCGGCTTAAACCGGTGAA	49163							
Qy	125	ACCTCTGCTCTACTATAAAACACAAAAAATTAGCCGGCGTGTGCGACGGCGCTGTAGTC	184							
Db	49164	ACCCGCTCTCTACTAAAAATACAAAAAATTAGCCGGCGCGCTGTAGTC	49223							
Qy	185	CCAGCTACTCGGAGGCTGAGGCAGGAGAAATGGCTTGAACCCGGGAGCGCGAACTTGAG	244							
Db	49224	CCAGCTACTTTGGAGGCTGAGGCAGGAGAAATGGCTTGAACCCGGGAGCGCGAGCTTGAG	49283							
Qy	245	TGAGCCGAGTTGCAACCACTGCACTCCAGCCCTGGGCAACACAGTGAGACTCCGCTCAAA	304							
Db	49284	TGAGCCGAGATTGCGCCACTGCACTCCAGCCCTGGGCGACAGAGCGAGACTCCGCTCAA	49343							
Qy	305	AAAAAAAAGAAAAGAAAGAAATCACA	333							
Db	49344	AAAAAAAAGAAAAGAAAGAAATCACA	49372							

```

RESULT 5
US-09-764-872-749
; Sequence 749, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 749
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-749

```

Query Match	27.8%;	Score 278.2;	DB 10;	Length 3407;
Best Local Similarity	91.3%;	Pred. No. 4.7e-60;		
Matches 295;	Conservative 0;	Mismatches 28;	Indels 0;	Gaps 0;
QY	6	AAGAAAGCACATCTCGCGCGGGCGTGGTCTCACACCTGTAAATCCCAACACTTTGGGAG	65	
Db	3068	AAAAAAGTGAAACACGCGCGGGCGGTGGCTCACGCTGTAAATCCCAACACTTTGGGAG	3127	
QY	66	GCCGAGGCGGGTGGATCATCTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGA	125	
Db	3128	GCCGAGGCGGGCGGATCATCGAGGTCAGGAGATCGAGACCATCTCTGGCTAACACGGTGA	3187	
QY	126	CCCTGTCTCTACTTAAAAACACAAAAATTTAGCGGGCGTGGTGGCAGGCGCTGTAGTCC	185	
Db	3188	CCCGCTCTCTACTTAAAAATACAAAAATTTAGCGGGCGTGGTGGCGGGCGCTGTAGTCC	3247	
QY	186	CAGCTACTCGGGAGGCTGAGGCAGGAGATGGCTGAACCCGGGAGGCGGAATTTGCAGT	245	
Db	3248	CAGCTACTCGGGAGGCTGAGGCAGGAGATGGCGTGAACCCGGGAGGCGGAGCTTGCAGT	3307	
QY	246	GAGCCGAGGTTGCACCACTGCACTCCAGCGCTGGGCAACACAGTGAGACTCCGCTCTCA	305	
Db	3308	GAGCCGAGTGCAGCCACTGCACTCCAGCGTGGGCAACAGGCGAGACTCCGCTCTCA	3367	
QY	306	AAAAAAGAAAGAAAAAGAAA	328	
Db	3368	AAAAAAGAAAGAAAAAGAAA	3390	

```

RESULT 6
US-10-148-806-3
; Sequence 3, Application US/10148806
; Publication No. US20030138933A1
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; TITLE OF INVENTION: HELICASE
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/148,806
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-148-806-3

```

Query Match	27.7%;	Score 276.8;	DB 15;	Length 114793;
Best Local Similarity	92.9%;	Pred. No. 5.2e-59;		
Matches 290;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0;
Qy 17	TCTCGCCGGGCGTGGTCTCACACCTGTGAATCCCAACACTTTGGGAGGCCGAGGCGGG	76		
Db 19284	TCTCGCCGGGCGGCTGGCTCACCCCTGTGAATCCAGCACTTTGGGAGGCCGAGGCGGG	19343		
Qy 77	TGGATCACTAGCTCAGGAGATCGGAGACCATCTGTGCTGACACCGGTGAACCCCTGCTCTA	136		
Db 19344	CGGATCAGAGTTCAGGAGATCGAGACCATCTGTGCTAACACCGGTGAACCCCTGCTCTA	19403		
Qy 137	CTTAAACAACAAAAATTAGCGGCGGTGTGGCAGGCGCTGTAGTCCCAAGCTACTCGG	196		
Db 19404	CTTAAAAATCAAAAAATTAGCGGCGGTGTGGCAGGCGCTGTAGTCCCAAGCTACTCGG	19463		
Qy 197	GAGGCTGAGCGAGGAGATGGCCTGAACCCGGGAGGCGGAACCTTGCAGTGAGCCGAGGTT	256		
Db 19464	GAGGCTGAGCGAGGAGATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATC	19523		
Qy 257	GCACCATCTCACTCCAGCGCTTGGGCAACACAGTGAAGACTCCCGTCTCAAAAAAAGAA	316		
Db 19524	GGCCACTGCACTCCAGCGCTTGGGCGACAGAGCGAGACTCCGCTCTCAAAAAAAGAA	19583		
Qy 317	AAAGAAAAGAAA	328		
Db 19584	AAAAAAGAAAAA	19595		

```

RESULT 7
US-10-859-792-3
; Sequence 3, Application US/10859792
; Publication No. US20050136425A1
; GENERAL INFORMATION:
; APPLICANT: Bai, Chang
; APPLICANT: Metzger, Michael
; APPLICANT: Liu, Xiaomei
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NHL, A DNA
; TITLE OF INVENTION: HELICASE
; FILE REFERENCE: 20585P
; CURRENT APPLICATION NUMBER: US/10/859,792
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/10/148,806
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US00/33065
; PRIOR FILING DATE: 2000-12-09

```



```
; PRIOR APPLICATION NUMBER: 60/169,970
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 114793
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-859-792-3

Query Match      27.7%; Score 276.8; DB 22; Length 114793;
Best Local Similarity 92.9%; Pred. No. 5.2e-59;
Matches 290; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 17 TCTCGCGCGGCGTGGCTCACACCTGTATATCCACACATTTTGGGAGGCGGAGCGGG 76
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19284 TCTCGCGCGGCGGCTGGCTCACACCTGTATATCCACACATTTTGGGAGGCGGAGCGGG 19343
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 77 TGGATCACTAGGTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGAACCCCTGTCTTA 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19344 CGGATCAGAGTCAGGATCGAGACCATCTCTGGCTGAACCGTGAACCCCTGTCTTA 19403
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 137 CTAACAAACAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19404 CTAACAAATACAAAAATTTAGCGGCGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGG 19463
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 197 GAGCTGAGGCGAGGATGCGCTGACCCGCGGAGCGGAACTTGCAGTGAGCGGAGTT 256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19464 GAGCTGAGGCGAGGATGCGCTGACCCGCGGAGCGGAGCTTGCAGTGAGCGGAGATC 19523
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 257 GCACCACTGCACTCAGGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAA 316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19524 GCGCACTGCACTCAGGCTGGGCGACAGACGAGACTCCGCTCTCAAAAAAAGAA 19583
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 317 AAACAAAAGAAA 328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19584 AAAAAAAGAAA 19595

RESULT 8
US-10-174-319-15/c
; Sequence 15, Application US/10174319
; Publication No. US20030232771A1
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Susan M. Preier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF MARK3 EXPRESSION
; FILE REFERENCE: PTS-0018
; CURRENT APPLICATION NUMBER: US/10/174,319
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 121
; SEQ ID NO 15
; LENGTH: 119501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-174-319-15

Query Match      27.6%; Score 276.6; DB 17; Length 119501;
Best Local Similarity 91.0%; Pred. No. 6e-59;
Matches 294; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 2 ACAGAGAAAGACATCTCGCGCGGCGTGGTGGCTACACCTGTAATCCCAACACTTTG 61
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67931 ACACAAAAGACAAATCTCGCGCGGCGTGGTGGCTACCGCTGTAATCCCAACACTTTG 67872
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 62 GGAGCCCGAGCGGTGATCACTAGTTCAGGAGATGAGACCACTCTGGCTGACACCGT 121
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67871 GGAGCCCGAGCGGTGATCACTAGTTCAGGAGATGAGACCACTCTGGCTGACACCGT 67812
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 122 GAAACCCCTGTCTACTAAAAACAAAAAATTTAGCCGGGCGTGGTGGCAGCGCGCTGTA 181
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



QY 306 AAAAAAAAAAGAGAAAGAAAGAAATCACATCTCATTTCAAGTGGTGG 350  
|||||  
Db 75310 AAAAAAAAAAGAGAAAGAAAGAAATCCAGGCGAGTTGAAACAGTAG 75266  
|||||

## RESULT 13

US-10-779-271-1/c  
; Sequence 1, Application US/10779271  
; Publication No. US20040220387A1  
; GENERAL INFORMATION:  
; APPLICANT: AstraZeneca AB  
; TITLE OF INVENTION: Methods  
; FILE REFERENCE: AS2D-P02-251  
; CURRENT APPLICATION NUMBER: US/10/779,271  
; CURRENT FILING DATE: 2004-02-13  
; PRIOR APPLICATION NUMBER: 09/463,844  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: PCT/GB98/00259  
; PRIOR FILING DATE: 1998-07-28  
; PRIOR APPLICATION NUMBER: 9716162.4  
; PRIOR FILING DATE: 1997-08-97  
; PRIOR APPLICATION NUMBER: 60/535,986  
; PRIOR FILING DATE: 2004-01-12  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 344805  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-779-271-1

Query Match 27.6%; Score 276; DB 20; Length 344805;  
Best Local Similarity 89.5%; Pred. No. 1.4e-58;  
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTACACCTGTAAATCCCAACACTTTG 61  
Db 88752 ACTTAAGAAATGGAATTCGCGCGCGGTGGTGTGCTACACCTGTAAATCCCAACACTTTG 88693  
|||||  
QY 62 GGAGCCCGAGGGTGTGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGT 121  
Db 88692 GGAGCCCGAGGGCGGGGATCACGAGGTTCAGGAGATGAGACCATCTCTGGCTGACACCGT 88633  
|||||  
QY 122 GAAACCCCTGTCTTACTAAAAACACAAAAATTAGCCGGCGTGTGTCAGCGCTGTGA 181  
Db 88632 GAAACCCCTGTCTTACTAAAAATACAAAAATTAGCCGGCGTGTGTCAGCGCTGTGA 88573  
|||||  
QY 182 GTCCAGCTACTCGGAGGCTGAGGAGGAGAAATGGCTGAACCCGGGAGCGGAACTTG 241  
Db 88572 GTCCAGCTACTTGGGAGGCTGAGGAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTG 88513  
|||||  
QY 242 CAGTGAGCCGAGGTGACACCTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCTC 301  
Db 88512 CAGTGAGCCGAGATTGCGCCACTGCACTCCAGCTGGGCGAGAGAGACTCCGTCTC 88453  
|||||  
QY 302 AAAAAAAAAAGAGAAAGAAAGAAATCACA 333  
Db 88452 AAAAAAAAAAGAGAAAGAAAGAAATCACA 88421  
|||||

## RESULT 14

US-10-737-082-70/c  
; Sequence 70, Application US/10737082  
; Publication No. US20050130170A1  
; GENERAL INFORMATION:  
; APPLICANT: Bayer Healthcare LLC  
; APPLICANT: Beard, Chris  
; APPLICANT: Burgess, Chris  
; APPLICANT: Gannon, Allison  
; APPLICANT: Harvey, Jeanne  
; APPLICANT: Lechner, John F.  
; APPLICANT: Li, Zheng  
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

; FILE REFERENCE: 1657/2032  
; CURRENT APPLICATION NUMBER: US/10/737,082  
; CURRENT FILING DATE: 2003-12-16  
; PRIOR APPLICATION NUMBER: US 10/737,082  
; PRIOR FILING DATE: 2003-12-16  
; NUMBER OF SEQ ID NOS: 300  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 70  
; LENGTH: 354592  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-737-082-70

Query Match 27.6%; Score 276; DB 22; Length 354592;  
Best Local Similarity 89.5%; Pred. No. 1.4e-58;  
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTCACACCTGTAAATCCCAACACTTTG 61  
Db 91143 ACTTAAGAAATGGAATTCGCGCGCGGTGGTGTGCTCACACCTGTAAATCCCAACACTTTG 91084  
|||||  
QY 62 GGAGCCCGAGGGTGGATCACTAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGT 121  
Db 91083 GGAGCCCGAGGGCGGGGATCACGAGTTCAGGAGATGAGACCATCTCTGGCTGACACGGT 91024  
|||||  
QY 122 GAAACCCCTGTCTTACTAAAAACACAAAAATTAGCCGGCGTGTGTCAGCGCTGTGA 181  
Db 91023 GAAACCCCTGTCTTACTAAAAATACAAAAATTAGCCGGCGTGTGTCAGCGCTGTGA 90964  
|||||  
QY 182 GTCCAGCTACTCGGAGGCTGAGGAGGAGAAATGGCTGAACCCGGGAGCGGAACTTG 241  
Db 90963 GTCCAGCTACTTGGGAGGCTGAGGAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTG 90904  
|||||  
QY 242 CAGTGAGCCGAGGTGACACCTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCTC 301  
Db 90903 CAGTGAGCCGAGATTGCGCCACTGCACTCCAGCTGGGCGAGAGAGACTCCGTCTC 90844  
|||||  
QY 302 AAAAAAAAAAGAGAAAGAAAGAAATCACA 333  
Db 90843 AAAAAAAAAAGAGAAAGAAAGAAATCACA 90812  
|||||

## RESULT 15

US-10-765-790-70/c  
; Sequence 70, Application US/10765790  
; Publication No. US20050130172A1  
; GENERAL INFORMATION:  
; APPLICANT: Bayer Healthcare LLC  
; APPLICANT: Beard, Chris  
; APPLICANT: Burgess, Chris  
; APPLICANT: Gannon, Allison  
; APPLICANT: Harvey, Jeanne  
; APPLICANT: Lechner, John F.  
; APPLICANT: Li, Zheng  
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
; FILE REFERENCE: 1657/2035  
; CURRENT APPLICATION NUMBER: US/10/765,790  
; CURRENT FILING DATE: 2004-01-27  
; PRIOR APPLICATION NUMBER: US 10/737,082  
; PRIOR FILING DATE: 2003-12-16  
; NUMBER OF SEQ ID NOS: 300  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 70  
; LENGTH: 354592  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-765-790-70

Query Match 27.6%; Score 276; DB 22; Length 354592;  
Best Local Similarity 89.5%; Pred. No. 1.4e-58;  
Matches 297; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
QY 2 ACAGAGAAAGCACATCTCGCGCGCGGTGGTGTGCTCACACCTGTAAATCCCAACACTTTG 61  
Db 91143 ACTTAAGAAATGGAATTCGCGCGCGGTGGTGTGCTCACACCTGTAAATCCCAACACTTTG 91084  
|||||

```
Db 91143 ACTTAAGAAATGGAATTCGGCGCGCGGTGGGTCAAGCCTGTAAATCCCGAGCACTTGG 91084
QY 62 GGAGGCCGAGCGCGGTGGATCACTAGGTCAGGAGATGGAGACCATCCTGGCTTGACACGGT 121
Db 91083 GGAGGCCGAGCGCGCGGTATCAGGAGTCAAGAGATCGAGACCATCCCGGCTAAAAACGGT 91024
QY 122 GAAACCCCTGTCTCTACTAAAAACACAAAAAATTAGCCGGCGGTGGTCAGGCGCCTGTA 181
Db 91023 GAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGCGGTAGTGGCGGCGCCTGTA 90964
QY 182 GTCCAGCTACTCGGAGGCTGAGCGAGAGATGGCCTGAACCCGGGAGGCGGAACCTTG 241
Db 90963 GTCCAGCTACTTGGAGGCTGAGCGAGAGATGGCGTGAACCCGGGAGGCGGAGCTTG 90904
QY 242 CAGTGAGCGAGTTGCAACCACTGCACTCCAGCCTGGGCAACACAGTGAGACTCCGTCTTC 301
Db 90903 CAGTGAGCGAGATTGCGCCTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTTC 90844
QY 302 AAAAAAAAAAAGAAAGAAAGAAATCACA 333
Db 90843 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 90812
```

Search completed: July 4, 2005, 02:13:07  
Job time : 406.727 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2005, 00:19:06 ; Search time 2176.56 Seconds

(without alignments)  
17505.766 Million cell updates/sec

Title: US-09-936-271C-13\_COPY\_8000\_9000

Perfect score: 1001

Sequence: 1 aacagagaagacacatctc.....taaaataactaatgttttt 1001

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl1:\*

9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282	28.2	822	4	BG249643 602319736
2	277	27.7	759	9	AG185133 Pan trogl
3	274.2	27.4	564	5	BU860291 AGENCOURT
4	274.2	27.4	566	5	BU957747 AGENCOURT
5	273.6	27.3	691	4	BF965007 602268753
6	273.2	27.3	444	8	B89781 CIT-HSP-217
7	272.6	27.2	745	9	AG029211 Pan trogl
8	272	27.2	1073	8	BZ601227 WHACR89TR
9	271.8	27.2	664	8	AQ343449 RPC111-12
10	271.8	27.2	834	8	AO744681 HS 5505 A
11	271	27.1	706	6	CA416075 UI-H-F80
12	270.4	27.0	709	9	AG167438 Pan trogl
13	270.2	27.0	517	1	AL046409 DKFp2434N
14	270	27.0	1051	8	BZ598731 WHACR86TR
15	269.8	27.0	731	8	AQ315280 RPC111-10
16	269.4	26.9	576	6	CA427630 UI-H-DFO
17	269	26.9	612	8	AQ395455 CITBI-EI-
18	268.8	26.9	375	2	AW303196 xtr89h02.x
19	268.8	26.9	513	5	BX502388 DKFp779M
20	268.6	26.9	669	9	AG155696 Pan trogl
21	268.6	26.8	337	1	AI284640 qu23h09.x
22	268.6	26.8	469	6	CB217725 NISC nb04
23	268.2	26.8	513	5	BX42735 DKFp781F
24	268.2	26.8	541	6	CA437967 UI-H-DHO-

C 25	268.2	26.8	630	8	AO628459	CITBI-EI-
C 26	268.2	26.8	715	5	BU617628	UI-H-DFO-
27	268	26.8	662	9	AG144106	Pan trogl
28	268	26.8	1050	6	CD252308	AGENCOURT
29	267.8	26.8	379	2	AW301350	xs75g06.x
30	267.8	26.8	514	8	AQ002001	CIT-HSP-2
31	267.8	26.8	672	9	AG173650	Pan trogl
32	267.6	26.7	515	8	AQ019249	CIT-HSP-2
33	267.4	26.7	657	9	AG182991	Pan trogl
34	267	26.7	585	4	EM479763	AGENCOURT
35	267	26.7	647	4	EM559057	AGENCOURT
C 36	267	26.7	701	5	BU615970	UI-H-DFO-
C 37	266.8	26.7	348	2	AW274349	xs74g09.x
38	266.8	26.7	400	5	BX482403	DKFp686G
39	266.8	26.7	581	2	BF677892	602084710
40	266.4	26.6	612	7	CR817709	IB35d10.Y
C 41	266.4	26.6	815	8	BZ608327	WHACR28TR
42	266.2	26.6	666	9	AG075258	Pan trogl
C 43	266	26.6	650	9	AG116352	Pan trogl
44	266	26.6	3680	3	HSM802780	Homo sapi
45	265.8	26.6	354	5	BU963331	AGENCOURT

## ALIGNMENTS

RESULT 1  
BG249643  
LOCUS  
DEFINITION BG249643 822 bp mRNA linear EST 13-FEB-2001  
602319736F1 NIH\_MGC\_89 Homo sapiens CDNA clone IMAGE:4414905 5',  
mRNA sequence.  
ACCESSION BG249643  
VERSION BG249643.1 GI:12759459  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (Bases 1 to 822)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM10143 row: h column: 10  
High quality sequence stop: 546.  
Location/Qualifiers  
1..822  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4414905"  
/tissue\_type="hypernephroma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_89"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## FEATURES

Source  
Query Match : 28.2%; Score 282; DB 4; Length 822;  
Best Local Similarity 87.3%; Pred No. 7.4e-32;  
Matches 309; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
ORIGIN

```
Qy 6 AAGAAAGCACATCTCGCGCGCGGTGGGTCTCACCCTGTAATCCCAACACTTTGGAG 65
Db 192 AAAATTCAGATATCGGCGCGCGGTGGGTCTCAGCGCTGTAATCCCAACACTTTGGAG 251
Qy 66 GCCGAGGCGGTGGATCACTAGGTACAGAGATGAGACCATCTCGGTGACACCGTGAAA 125
Db 252 GCCGAGGCGCGCGGTGGATCACTAGGTACAGAGATGAGACCATCTCGGTGTAACACCGTGAAA 311
Qy 126 CCCTGTCTCTACTATAAAACACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCC 185
Db 312 CCCCCTCTCTACTATAAAATACAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCC 371
Qy 186 CAGTACTCGGGAGGCTGAGGAGAGAAATGGCTGAAACCCGGAGGCGGAACTTGCAGT 245
Db 372 CAGTACTCGGGAGGCTGAGGAGAGAAATGGCTGAAACCCGGAGGCGGAGCTTGCAGT 431
Qy 246 GAGCGAGGTTGCACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCCGTCTCAAAA 305
Db 432 GAGCGAGATGGCGCACTGCACCTCCAGCTGGGCGACAGAGCGAGACTCCGTCTCAAAA 491
Qy 306 AAAAAAAGAAAAAGAAAAAATCACATCTCATTTCAAGTGTGGCATTTAAAA 359
Db 492 AAAAAAAGAAAAAACAACAAAAAATAAAAAAAGGGGGGGGCGCAAAAA 545

RESULT 2
AG185133 759 bp DNA linear GSS 09-JAN-2002
LOCUS Pan troglodytes DNA, clone: RP43-059C02.T7, genomic survey
DEFINITION
ACCESSION AG185133
VERSION AG185133.1 GI:16714813
KEYWORDS GSS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
1 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tokoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library RPCI-43
2 (bases 1 to 759)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Tokoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbese@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T7
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. 759
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-059C02.T7"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN
Query Match 27.7%; Score 277; DB 9; Length 759;
Best Local Similarity 89.5%; Pred. No. 4.1e-31;
```

```
Matches 298; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Qy 14 ACATCTCGCGCGCGGTGGGTCTCACCCTGTAATCCCAACACTTTGGAGGCGGAGGC 73
Db 286 AAAAATCGCGCGCGGTGGGTCTCAGCTGTAATCCAGCACTTTGGAGGCGGAGGC 345
Qy 74 GGTGGATCACTAGGTACAGAGATGGAGACCATCTCGGTGACACCGTGAACCCCTGTCT 133
Db 346 GGGCGGATCAGAGTTCAGGATCGAGACCATCTCGGTGAACACCGGTGAACCCCGTCT 405
Qy 134 CTACTAAAAACACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCCAGCTACT 193
Db 406 CTACTAAAAATACAAAAAATTAGCCGGCGGTGGGTGGCGCGCTGTAGTCCAGCTACT 465
Qy 194 CGGAGGCTGAGGAGAGAAATGGCTGAAACCCGGAGGCGGAACTTGCAGTGAGCCGAG 253
Db 466 CGGAGGCTGAGGAGAGAAATGGCTGAAACCCGGAGGCGGAGCTTGCAGTGAGCCGAG 525
Qy 254 GTTGCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGTCTCAAAAAAATAAAA 313
Db 526 GTTGTCCTTCCTGCACTCCAGCTGGGCGACAGAGTGAGACTCCGTCTCAAAAAAATAAAA 585
Qy 314 AGAAAAAGAAAGAAATCACATCTCATTTCAAGTG 346
Db 586 AAAAAATCAAAATAGTATGTAAGTGGGTG 618

RESULT 3
BU860291/c 564 bp mRNA linear EST 16-OCT-2002
LOCUS AGENCOURT 10440048 NIH MGC 107 Homo sapiens cDNA clone
DEFINITION IMAGE:6651369 5', mRNA sequence.
ACCESSION BU860291
VERSION BU860291.1 GI:24045283
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 564)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2898 row: b column: 09
High quality sequence stop: 543.
FEATURES
Location/Qualifiers
1. 564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6651369"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
```



Query Match 27.4%; Score 274.2; DB 5; Length 564;  
 Best Local Similarity 91.2%; Pred. No. 1.2e-30;  
 Matches 291; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 17 TCTGGCGGGCGTGGTGGCTCACACCTGTATATCCAAACACTTTGGGAGCGCGGCGG 76  
 Db 328 TTTGGCGGGCGGCGTGGCTCACGCTGTATATCCAGCACTTTGGGAGCGCGGCGG 269

QY 77 TGGATCACTAGTCTAGGAGATGGAGACCATCTGTGCTGACACGCTGAAACCCCTGTCTA 136  
 Db 268 TGGATCACTAGTCTAGGAGATGGAGACCATCTGTGCTGACACGCTGAAACCCCTGTCTA 209

QY 137 CTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTCGG 196  
 Db 208 CTAAAAATACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTCGG 149

QY 197 GAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTTCAGTGAGCGGAGTT 256  
 Db 148 GAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAGCTTCAGTGAGCGGAGATC 89

QY 257 GCACCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAAAGA 316  
 Db 88 GCGCCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAAAGA 29

QY 317 AAAGAAAAGAAATCATCATC 335  
 Db 28 AAAAACTCGAGACTAGTTC 10

RESULT 4  
 BU957747/c  
 LOCUS  
 DEFINITION AGENCOURT\_10620302 NIH\_MGC\_107 Homo sapiens cDNA clone  
 IMAGE:6731834 5', mRNA sequence.

ACCESSION BU957747  
 VERSION BU957747.1 GI:24187319  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 566)  
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 TITLE Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LCM3058 row: 0 column: 01  
 High quality sequence stop: 450.  
 Location/Qualifiers  
 1..566  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6731834"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 107"  
 /note="Organ: breast; vector: pOTB7; Site: 1: EcoRI;  
 Site: 2: XhoI; cDNA made by oligo-dt priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

FEATURES  
source

FEATURES  
 source  
 1..566  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6731834"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 107"  
 /note="Organ: breast; vector: pOTB7; Site: 1: EcoRI;  
 Site: 2: XhoI; cDNA made by oligo-dt priming.  
 Directionally cloned into EcoRI/XhoI sites using the  
 following 5' adaptor: GGCACGAG(G). Library constructed by  
 Ling Hong in the laboratory of Gerald M. Rubin (University  
 of California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 27.4%; Score 274.2; DB 5; Length 566;  
 Best Local Similarity 91.2%; Pred. No. 1.2e-30;  
 Matches 291; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 17 TCTGGCGGGCGTGGTGGCTCACACCTGTATATCCAAACACTTTGGGAGCGCGGCGG 76  
 Db 328 TTTGGCGGGCGGCGTGGCTCACGCTGTATATCCAGCACTTTGGGAGCGCGGCGG 269

QY 77 TGGATCACTAGTCTAGGAGATGGAGACCATCTGTGCTGACACGCTGAAACCCCTGTCTA 136  
 Db 268 TGGATCACTAGTCTAGGAGATGGAGACCATCTGTGCTGACACGCTGAAACCCCTGTCTA 209

QY 137 CTAAAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTCGG 196  
 Db 208 CTAAAAATACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCCCTGTAGTCCAGCTACTCGG 149

QY 197 GAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAACTTCAGTGAGCGGAGTT 256  
 Db 148 GAGGCTGAGGCGAGGAGATGGCTGAACCCCGGAGCGGAGCTTCAGTGAGCGGAGATC 89

QY 257 GCACCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAAAGA 316  
 Db 88 GCGCCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAAAGAAAAAGA 29

QY 317 AAAGAAAAGAAATCATCATC 335  
 Db 28 AAAAACTCGAGACTAGTTC 10

RESULT 5  
BU965007

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```
ORIGIN
Query Match      27.3%; Score 273.6; DB 4; Length 691;
Best Local Similarity 89.6%; Pred. No. 1.3e-30;
Matches 294; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 1 AACAGAGAACACATCTCGCGCGGCGTGGCTCACACCTGTAATCCCAACACTTT 60
Db 132 AAAAGTTGAAGAAAAGATCGCGCGGCGGCTCACCCCTGTAATCCAGACTTT 191
Qy 61 GGGAGGCGAGCGCGGTGGATCACTAGGTCAAGGATGAGACCATCTCTGGCTGACACGG 120
Db 192 GGGAGGCAAGCGGGTGGATCAAGGTCAAGGATGAGACCATCTCTGGCTAACACGG 251
Qy 121 TGAACCTGTCTCTACTATAAACAACAAAAATTAGCGCGGCGTGGTGGCAGGCGCTGT 180
Db 252 TGAACCTGTCTCTACTATAAATAACAAAAATTAGCGCGGCGTGGTGGCAGGCTGT 311
Qy 181 AGTCCAGCTACTCGGAGGCTGAGGACGAGGATGGCTGAACCGGAGCGGAACTT 240
Db 312 AGTCCAGCTACTCGGAGGCTGAGGACGAGGATGGCTGAACCGGAGCGGAGAGCTT 371
Qy 241 GCAGTGAGCGAGGTTGCACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCCGCT 300
Db 372 GCAGTGAGCGAGGTTGGCCACTGCACTCCAGCTGGGTCACAGTGAGACTCCGCTCC 431
Qy 301 CAAAAAAGAAAAAGAAAAAGAAAAA 328
Db 432 CAAAAAAGAAAAAACAACCAACAAAA 459

RESULT 6
B89781
LOCUS CIT-HSP-2171J2.TR CIT-HSP Homo sapiens genomic clone 2171J2,
genomic survey sequence.
ACCESSION B89781
VERSION B89781.1 GI:2972261
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Other GSSs: CIT-HSP-2171J2.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
1. 444
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7103508"
/db_xref="taxon:9606"

FEATURES
source
1. 444
```

```
ORIGIN
Query Match      27.3%; Score 273.2; DB 8; Length 444;
Best Local Similarity 90.6%; Pred. No. 1.8e-30;
Matches 290; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 9 AAAGCACATCTCGCGCGGCGTGGCTCACACTGTAATCCCAACACTTTGGAGGCC 68
Db 60 AAAGCTCAATAGCGCGGCGGCTGCTCAGCCTGTAATCCCAAGCACTTTGGAGGCC 119
Qy 69 GAGCGGGTGGATCAGTGTAGGAGATGAGAGACCATCTGGCTGACACGTTGAAACCC 128
Db 120 GAGCGGGTGGATCAGGAGTCAAGGATGAGAGACCATCTGGCTTAACACGTTGAAACCC 179
Qy 129 TGTCTCTACTAAAAACACAAAAAATTAGCGCGGCGTGGTGGCAGCGCTGTAGTCCCAG 188
Db 180 CGTCTCTACTAAAAATACAAAAATTAGCGCGGCGTGGTGGCAGCGCTGTAGTCCCAG 239
Qy 189 CTACTCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGGAACTTGCAGTGAG 248
Db 240 CTACTCGGAGGCTGAGCGAGGAGAAATGGCTGAACCCGGGAGCGGAGCTTGCAGTGAG 299
Qy 249 CCGAGGTTGACCACTGCACCTCCAGCTGGGCAACACAGTGAGACTCGTCTCAAAAAA 308
Db 300 CCGAGATCAGCCCACTGCACCTCCAGCTGGGCGAGAGAGAGACTCGTCTCAAAAAA 359
Qy 309 AAAAAAGAAAAAGAAAAA 328
Db 360 AAAAAAGAAAAAGAAAAA 379

RESULT 7
AG029211/c
LOCUS AG029211 Pan troglodytes DNA, clone: PTB-001D16.F, genomic survey sequence.
DEFINITION AG029211
ACCESSION AG029211
VERSION AG029211.1 GI:16556083
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 745)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
(E-mail:chimpbes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 745

FEATURES
source
```

```

/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-001D16.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

```

## ORIGIN

```

Query Match      27.2%; Score 272.6; DB 9; Length 745;
Best Local Similarity 89.6%; Pred. No. 1.8e-30;
Matches 293; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 2 ACAGAAAGAACACATCTCGCGCGCGCTGGTGCATCAGCTGAATCCCAACACTTTG 61
Db 415 ATAGAAAGAGCTCCTCGCGCGCGCGCTGGTGCATCAGCTGAATCCCAACACTTTG 356
QY 62 GGAGCGGAGCGGTGGATCACTAGTTCAGGAGATGGAGACCATCTCTGGCTGACACGGT 121
Db 355 GGAGCGGAGCGGTGGATCACTAGTTCAGGAGATGGAGACCATCTCTGGCTGACACGGT 296
QY 122 GAAACCCCTGTCTTACTTAAACACAAAAATTAGCCGGCGTGGTGCAGCGCGCTGTA 181
Db 295 GAAACCCCTGTCTTACTTAAACACAAAAATTAGCCGGCGTGGTGCAGCGCGCTGTA 236
QY 182 GTCCAGCTACTCGGAGGCTGAGCAGGAGATGGCTGAACCGGCGGAGCGGAACTTG 241
Db 235 GTCCAGCTACTCGGAGGCTGAGCAGGAGATGGCTGAACCGTGGGAGCGGAGCTTG 176
QY 242 CAGTGAGCGGAGGTTGCACCTGACCTCCAGCTGGGCAACACAGTCAGACTCCGCTCTC 301
Db 175 CAGTGAGCGGAGATCGCCCTGCTGCTCCAGCTGGGCGAGCAGAGCGAGCTCGCTCTC 116
QY 302 AAAAAAAGAAAAAGAAAAAGAAAA 328
Db 115 AAAAAAAGAAAAAGAAAAAGAAAAAGAA 89

```

## RESULT 8

```

BZ601227
LOCUS WHADC9TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
DEFINITION sapiens genomic clone MCF7_1-19P9, genomic survey sequence.
ACCESSION BZ601227
VERSION BZ601227.1 GI:31509689
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
Location/Qualifiers
1..1073
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

## FEATURES

## source

```

/clone="MCF7_1-19P9"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7.1)"
/note="vector: pECBAC1; Site 1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."

```

## ORIGIN

```

Query Match      27.2%; Score 272; DB 8; Length 1073;
Best Local Similarity 89.3%; Pred. No. 1.9e-30;
Matches 293; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 6 AAGAAAGACATCTCGCGCGCGCTGGTGCATCAGCTGAATCCCAACACTTTGGAG 65
Db 318 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 377
QY 66 GCCGAGCGGCTGATCACTAGTTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGA 125
Db 378 GCCGAGCGGCTGATCACTAGTTCAGGAGATGGAGACCATCTCTGGCTGACACGGTGA 437
QY 126 CCTGTCTTACTTAAACACAAAAATTAGCCGGCGTGGTGCAGCGCTGTAGTCC 185
Db 438 CCTGTCTTACTTAAACACAAAAATTAGCCGGCGTGGTGCAGCGCTGTAGTCC 497
QY 186 CAGTACTCGGAGGCTGAGCAGGAGATGGCTGAACCGGAGCGGAACTTGCACT 245
Db 498 CAGTACTCGGAGGCTGAGCAGGAGATGGCTGAACCGGAGCGGAACTTGCACT 557
QY 246 GAGCGGAGGTTGCACCTGACCTCCAGCTGGGCAACACAGTCAGACTCCGCTCAAAA 305
Db 558 GAGCGGAGATCCCGCCACTGCCTCCAGCTGGGCGAGCAGAGCGAGACTCCGCTCAAAA 617
QY 306 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 333
Db 618 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAA 645

```

## RESULT 9

```

BZ601227
LOCUS AQ343449
DEFINITION RPCI11-122K4.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-122K4,
genomic survey sequence.
ACCESSION AQ343449
VERSION AQ343449.1 GI:4168345
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers

```

## FEATURES

```

source
1. .664
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7546707"
/db_xref="taxon:9606"
/clones="RPCI-11-122K4"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
RPC11 Human Male BAC Library"

ORIGIN
Query Match 27.2%; Score 271.8; DB 8; Length 664;
Best Local Similarity 90.1%; Pred. No. 2.5e-30;
Matches 291; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 21 GCGCGGGCGGTGGCTCACACCTGTATATCCACACATTGGAGGCGGCGGTGGA 80
Db 42 GCGCGGGCGGTGGCTCACACCTGTATATCCACACATTGGAGGCGGCGGCGGA 101
Qy 81 TCACGTAGTCAGGAGATGAGACCATCTGGCTGACAGGTGAACCTGTCTACTAA 140
Db 102 TCACGTAGTCAGGAGATGAGACCATCTGGCTGACAGGTGAACCTGTCTACTAA 161
Qy 141 AAACACAAAAAATTAGCGCGGTGGTGGCAGGCGCTGTAGTCCAGCTACTCGGAGG 200
Db 162 AAATACAAAAAATTAGCAGCGGTGGTAGCGGCGCTGTAGTCCAGCTACTCGGAGG 221
Qy 201 CTGAGGCGAGGAGATGGCTGAACCCGGGCGGGAATTCAGTGAGCCGAGGTTGCAC 260
Db 222 CTGAGGCGAGGAGATGGCTGAACCCGGGCGGAGCTTCAGTGAGCCGAGATCGCGC 281
Qy 261 CACTGCATCTCAGCTGGGCAACACAGTGAATCCGCTCAAAAAAAGAAAG 320
Db 282 CACTGCATCTCAGCTGGGCAACACAGTGAATCCGCTCAAAAAAAGAAAG 341
Qy 321 AAAAGAAATCATCTCAATCAA 343
Db 342 AAAAAGGCTGAGTTGTTCAA 364

RESULT 10
AQ744681/c
LOCUS
DEFINITION HS_5505_A2_B04_SP6 RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1081 Col=8 Row=C, genomic survey sequence.
ACCESSION AQ744681
VERSION AQ744681.1 GI:5522203
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 834)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL 99380589
MEDLINE 10449764
PUBMED
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from

```

```

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1081 row: C column: 8
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 834.
Location/Qualifiers
source
1. .834
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=1081 Col=8 Row=C"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match 27.2%; Score 271.8; DB 8; Length 834;
Best Local Similarity 88.8%; Pred. No. 2.3e-30;
Matches 294; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 9 AAAGCACATCTCGCGCGGTGGTGGCTCACACCTGTAAATCCCAACACATTGGAGGCC 68
Db 544 AAATCAACATAGCGCGGTGGTGGCTCATGCTCTAATCCAGCACATTGGAGGCT 485
Qy 69 GAGCGGGTGGATCAGTGTAGGTGAGGAGATGAGACCATCTCTGGCTGACACGGTCAAAACCC 128
Db 484 GAGCGGGCGATCAGCGGTGAGGAGATGAGACCATCTCTGGCTGACACGGTCAAGGCC 425
Qy 129 TGCTCTTACTAAACACAAAAAATTAGCCGGGCGTGGTGGCAGCGCTGTAGTCCAG 188
Db 424 CGTCTCTACTAAAAATGCAAAAAAATTAGCCGGGCGAGGTGGTGGCGCTGTAGTCCAG 365
Qy 189 CTACTCGGAGGCTGAGGAGGAGATGGCTGAAACCCGGGCGGAGCTTGCAGTGTAG 248
Db 364 CTACTTGGGAGGCTGAGGAGGAGATGGCTGAAACCCGGGCGGAGCTTGCAGTGTAG 305
Qy 249 CCGAGGTTGCACCACTCACTCCAGCTGGGCAACACAGTCAAGTCCGTCTCAAAAAAA 308
Db 304 CCGAGATCGCGCCACTGCACTCCAGCTGGGCAACACAGTCAAGTCCGTCTCAAAAAAA 245
Qy 309 AAAAAAGAAAGAAAGAAATCAATCTCAT 339
Db 244 AAAAAAAGAAAGAAAGAAATCAATCTCAT 214

RESULT 11
CA416075/c
LOCUS
DEFINITION UI-H-PEO-bbs-o-16-0-UI.s1 NCI CGAP PE0 Homo sapiens cDNA clone
UI-H-PEO-bbs-o-16-0-UI 3', mRNA sequence.
ACCESSION CA416075
VERSION CA416075.1 GI:24778726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

```

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 The following repetitive elements were found in this cDNA  
 sequence: 11-299, >ALU (matched complement) 217-326, >ALU (matched  
 complement) 591-706, >THE1B/LTR/MaLR (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
 1..706  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-FEO-bbs-o-16-0-UI"  
 /tissue\_type="Chondrosarcoma Cell line"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP FEO"  
 /note="Vector: pTTT3-Pac (Pharmacia) with a modified  
 polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_FEO is  
 a cDNA library containing the following tissue(s): a pool  
 of 3 chondrosarcoma cell lines (grade 2) The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pTTT3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 CGCTACGGAC. The cell lines was provided by Dr James Martin  
 of University of Iowa.  
 TAG\_TISSUE=Human grade 2 chondrosarcoma cell line pool  
 TAG\_LIB=UI-H-FEO  
 TAG\_SEQ=CGCTACGGAC"

## ORIGIN

Query Match 27.1%; Score 271; DB 6; Length 706;  
 Best Local Similarity 90.6%; Pred. No. 3.2e-30;  
 Matches 289; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 AACAGAGAACACATCTCGCGCGCGGTGGTGGCTCACACCTGTAATCCCACTTT 60  
 Db |||||  
 Qy 319 AAGAAAGAAATAGTAAATCGCGCGCGGTGGTGGCTCACGCTGTAATCCGACACTT 260  
 Db |||||  
 Qy 61 GGGAGCGCGAGCGGGTGGATCAGTGGTGGAGATGGAGACCATCTGGCTGACACGG 120  
 Db |||||  
 Qy 259 GGGAGCGCGAGCGCGGTGACAGGTGAGAGATCGAGACCATCTGGCTAACACGG 200  
 Db |||||  
 Qy 121 TGAACCTGTCTTCTACTAAACACAAAAAATTAGCCGGCGGTGGTGGCAGGCGCTGT 180  
 Db |||||  
 Qy 199 TGAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGCGGTAGTGGCGGCGCTGT 140  
 Db |||||  
 Qy 181 AGTCCAGCTACTCGGGAGGTGAGGAGGAGGAGATGGCTGACCGGAGCGGAACTT 240  
 Db |||||  
 Qy 139 AGTCCAGCTACTCGGGAGGTGAGGAGGAGGAGATGGCTGAAACCCGAGAGCGGAGCTT 80  
 Db |||||  
 Qy 241 GCAGTGAGCGAGGTTGCACACTGCATCTCAGCTGGGCAACACAGTGAGACTCCGCT 300  
 Db |||||  
 Qy 79 GCAGTGAGCCAGATGTCGCACTGCATCTCCAGCTGGCGGACAGCGAGACTCCGCT 20  
 Db |||||  
 Qy 301 CAAAAAAGAAAAAGAAAA 319  
 Db |||||  
 Qy 19 CAAAAAAGAAAAAGAAAA 1

## RESULT 12

AG167438/c

LOCUS

DEFINITION

ACCESSION

AG167438 709 bp DNA linear GSS 09-JAN-2002  
 Pan troglodytes DNA, clone: RP43-035103.TJ, genomic survey  
 sequence.  
 AG167438

## VERSION

AG167438.1 GI:16697116

## KEYWORDS

GSS.

## SOURCE

Pan troglodytes (chimpanzee)

## ORGANISM

Pan troglodytes

## REFERENCE

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
 Totoki, Y., Watanabe, H. and Sakaki, Y.  
 BAC end sequences of Library RPCI-43

## AUTHORS

2 (bases 1 to 709)  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
 Totoki, Y., Watanabe, H. and Sakaki, Y.

## TITLE

Unpublished  
 Direct Submission

## JOURNAL

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
 1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 (E-mail: chimpanzee@gsc.riken.go.jp URL: http://hgp.gsc.riken.go.jp/  
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
 Clones are derived from the chimpanzee BAC library RPCI-43 This BAC  
 end was generated during the R&D process and may have higher chance  
 of clone tracking errors.

## COMMENT

PRIMERS  
 Sequencing: TJ

## LIBRARY

Vector : pBACe3.6  
 R.Site 1 : EcoRI  
 R.Site 2 : EcoRI

## FEATURES

Location/Qualifiers  
 1..709

/organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="RP43-035103.TJ"  
 /sex="male"  
 /cell\_type="lymphocytes"  
 /clone\_lib="RPCI-43 Chimpanzee Male BAC Library"

## ORIGIN

Query Match 27.0%; Score 270.4; DB 9; Length 709;  
 Best Local Similarity 89.0%; Pred. No. 3.9e-30;  
 Matches 292; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 9 AAAGCACATCTCGCGCGCGGTGGTGGCTCACACCTGTAATCCCACTTTGGAGGCC 68  
 Db 578 AAAACTCTTCTGGCGCGGCACATGGCTCAGCGCTGTAATCCAGACTTTGGAGGCC 519  
 Qy 69 GAGCGGGTGGATCAGTGGTGGAGATGGAGACCATCTGGCTGACACCGTGAACCC 128  
 Db 518 AAGCGGGCGGATCAGCGGTGAGGATCAAGACCATCTCTGGCTAAACCGTGAACCC 459  
 Qy 129 TGTCTCTACTAAAAACAAAAAATTAGCCGGCGGTGGTGGCAGCGCTGTAGTCCCGAG 188  
 Db 458 TGTCTCTACTAAAAATACAAAAAATTAGTGGCGGTAGTGGCGGCGCTGTAGTCCCGAG 399  
 Qy 189 CTACTCGGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGGCGGAACTTGCAGTGAG 248  
 Db 398 CTACTCGGGAGGCTGAGCGAGGAGATGGCTGAAACCCGGGAGGCGGAGCTTGCAGTGAG 339  
 Qy 249 CCGAGGTTGACCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAAAA 308  
 Db 338 CCGAGATGGCGCCACTGCACCTCCAGCTGGGCGAGAGAGAGACTCCGCTCTCAAAAAA 279  
 Qy 309 AAAAAAGAAAAAGAAATACATCT 336  
 Db 278 AAAAAACAAAAAATACTGCTTCT 251

## RESULT 13

AL046409/c

LOCUS

DEFINITION

AL046409 517 bp mRNA linear EST 04-SEP-2003  
 DKEPz434N147\_s1 434 (synonym: htes3) Homo sapiens cDNA clone

DKFP2p434N147 3', mRNA sequence.  
AL046409  
VERSION AL046409.1 GI:5434485  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 517)  
AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE EST (Koehrer, et al.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 3' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BPFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.  
r1 sequence also available.  
This clone (DKFP2p434N147) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
FEATURES  
source  
1..517  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFP2p434N147"  
/tissue="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/lab\_lib="434 (synonym: htes3)"  
/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

ORIGIN  
Query Match 27.0%; Score 270.2; DB 1; Length 517;  
Best Local Similarity 89.8%; Pred. No. 4.6e-30;  
Matches 290; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
Qy 6 AAGAAGCACATCTCGGCGGGCGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 65  
Db 325 AAAAAAGTATTATAGGCGGGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 266  
Qy 66 GCCGAGCGGGTGTATCTAGTGTGAGAGATGAGACCATCTCTGGCTGACACGGTGAAA 125  
Db 265 GCCGAGCGGGCGGATCAGAGGTGAGAGATGAGACCATCTCTGGCTGAGAGTAAAACGGTGAAA 206  
Qy 126 CCCTGTCTCTACTAAAAACACAAAAATTAGCCGGCGTGGTGGCAGCGCCTGTAGTCC 185  
Db 205 CCGCGTCTCTACTAAAAATACAAAAATAGTGGCGGTGGTGGCGGCGCCTGTAGTCC 146  
Qy 186 CAGTCTCTCGGAGGCTGAGGAGAGATGGCTGAAACCCGGAGGCGGAACCTTGCACT 245  
Db 145 CAGTCTCTCGGAGGCTGAGGAGAGATGGCTGAAACCCGGAGGCGGAGCTTGCACT 86  
Qy 246 GAGCGAGGTTGCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAA 305  
Db 85 GAGCGAGATCCCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCCGCTCTCAAAC 26  
Qy 306 AAAAAAAGAAAAAGAAAAAGAAA 328  
Db 25 AAAAAAAGAAAAAAGAAAAA 3

RESULT 14  
BZ598731  
LOCUS BZ598731 1051 bp DNA linear GSS 08-JUN-2003  
DEFINITION WHACR86TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo sapiens genomic clone MCF7\_1-1704, genomic survey sequence.  
ACCESSION BZ598731

BZ598731.1 GI:31507193  
GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1051)  
AUTHORS Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J.W. and Collins, C.  
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
MEDLINE 22709111  
PUBMED 12788976  
COMMENT Contact: Volik SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
http://www.genomex.com  
Class: BAC ends.  
FEATURES  
source  
1..1051  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-1704"  
/sex="female"  
/clone\_lib="Human MCF7 breast cancer cell line library (MCF7\_1)"  
/note="Vector: pECBAC1; Site\_1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN  
Query Match 27.0%; Score 270; DB 8; Length 1051;  
Best Local Similarity 86.8%; Pred. No. 3.8e-30;  
Matches 297; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
Qy 6 AAGAAGCACATCTCGGCGGGCGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 65  
Db 226 AAAAAAGTATTATAGGCGGGCGGTGGTGTACACCTGTAAATCCCAACACTTTGGAG 285  
Qy 66 GCCGAGCGGGTGGATCCTAGGTGAGAGATGAGACCATCTCTGGCTGACACGGTGAAA 125  
Db 286 GCTGAGGCGAGTGGATCAGGAGGTGAGAGAACGAGACCATCTCTGGCTTAACACGGTGAAA 345  
Qy 126 CCCTGTCTCTACTAAAAACACAAAAATTAGCCGGCGTGGTGGCAGCGCCTGTAGTCC 185  
Db 346 CCGCGTCTCTACTAAAAATACAAAAATAGTGGCGGTGGTGGCGGCGCCTGTAGTCC 405  
Qy 186 CAGTCTCTCGGAGGCTGAGGAGAGATGGCTGAAACCCGGAGGCGGAACCTTGCACT 245  
Db 406 CAGTCTCTCGGAGGCTGAGGAGAGATGGCTGAGCCAGGAGCGGAGCTTGCACT 465  
Qy 246 GAGCGAGGTTGCACTGCACTCCAGCTGGGCAACACAGTGAGACTCCGCTCTCAAAA 305  
Db 466 GAGCGAGATCCCGCACTGCACTCCAGCTGGGCAACAGAGTGAAGTCCGCTCAAAAAA 525  
Qy 306 AAAAAAAGAAAAAGAAAAAGAAATCAGATCTCATTCAGTGG 347  
Db 526 AAAAAAAGAAAAAAGAAAAAAGAAATCAGATCTATGATTTG 567

RESULT 15  
AQ315280  
LOCUS AQ315280 731 bp DNA linear GSS 04-MAY-1999  
DEFINITION RPC111-104N24.TV RPC1-11 Homo sapiens genomic clone RPC1-11-104N24, genomic survey sequence.

ACCESSION AQ315280  
VERSION AQ315280.1 GI:4046743  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 731)  
Adams M.D., Rounseley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,  
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.  
Use of human BAC End Sequences for Sequence-Ready Map Building  
Unpublished (1998)  
Other GSSs: RPC111-104N24.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbs@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@jomb.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: 17  
Class: BAC ends.  
Location/Qualifiers  
1..731  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7539887"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-104N24"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC111 Human Male BAC Library"

## ORIGIN

Query Match 27.0%; Score 269.8; DB 8; Length 731;  
Best Local Similarity 90.0%; Pred. No. 4.7e-30;  
Matches 289; Conservative 0; Mismatches 32; Indels 0; Gaps 0;  
Qy 6 AAGAAAGCACATCTCGGCGGCGGTGGTGGCTCACACCTGTAATCCCAACACTTTGGGAG 65  
Db |||||  
124 AAGAAAGAGGTTAGGCGGCGGCGGTGGCTCACGCTGTATATCCAGCACTTTGGGAG 183  
Qy 66 GCCGAGCGGTGATCATCTAGGTGAGGATGGAGACCATCTCGGTGACACCGTGAAA 125  
Db |||||  
184 GCCGAGCGGCGGATCACGAGGTGAGGATCGAGACCATCCCGGCTAAACCGTGAAA 243  
Qy 126 CCCTGTCTCTACTAAAAACAAAAAATTAGCCGGCGGTGGTGGCGGCGCTGTAGTCC 185  
Db |||||  
244 CCCCCTCTCTACTAAAAATACAAAAAATTAGCCGGCGGTAGTGGCGGCGCTGTAGTCC 303  
Qy 186 CAGTACTCTCGGAGGCTCAGGCGAGGAGATGGCCCTGAACCCGGGAGGGGAACTTGCAGT 245  
Db |||||  
304 CAGTACTCTGGAGGCTCAGGCGAGGAGATGGCGTGAACCCGGGAGGTGGAGCTTGCAGT 363  
Qy 246 GAGCGAGGTTGCACCTGCACTCCAGCTGGGCAACACAGTGAAGACTCCGCTCTCAAA 305  
Db |||||  
364 GAGCGAGATCCCGCACTGCACTCCAGCTGGGCGACAGAGCGAGACTCCGCTCTCAAA 423  
Qy 306 AAAAAAAGAAAAAGA 326  
Db |||||  
424 AAAAAAAGAAAAAGA 444

Search completed: July 4, 2005, 15:07:14  
Job time : 2183.56 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 15:07:24 ; Search time 73 Seconds  
(without alignments)  
1552.340 Million cell updates/sec

Title: US-09-936-271c-14

Perfect score: 1608

Sequence: 1 MATARPPMWVLCALITALL.....VVTNLCKFTKWIQTIQANS 293

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq\_16Dec04:\*

1: geneseqp1980s:\*

2: geneseqp1990s:\*

3: geneseqp2000s:\*

4: geneseqp2001s:\*

5: geneseqp2002s:\*

6: geneseqp2003as:\*

7: geneseqp2003bs:\*

8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	293	2	Aay16777 Human ker
2	1608	100.0	293	2	Aay30524 Human PDS
3	1608	100.0	293	2	Aay38412 Human sec
4	1608	100.0	293	2	Aay38426 Human sec
5	1608	100.0	293	3	Aay66726 Membrane-
6	1608	100.0	293	3	Aab21296 Human KLIK
7	1608	100.0	293	4	Aau12399 Human PRO
8	1608	100.0	293	4	Aab65249 Human PRO
9	1608	100.0	293	5	Aau81966 Human PRO
10	1608	100.0	293	6	Abu58064 Human PRO
11	1608	100.0	293	6	Abu59142 Novel hum
12	1608	100.0	293	6	Abu82654 Human sec
13	1608	100.0	293	6	Abu17843 Novel hum
14	1608	100.0	293	6	Ada57425 Human sec
15	1608	100.0	293	6	Ada56974 Human sec
16	1608	100.0	293	6	Ada57427 Human sec
17	1608	100.0	293	6	Ada57428 Human sec
18	1608	100.0	293	6	Abu60573 Human sec
19	1608	100.0	293	6	Abu13955 Human PRO
20	1608	100.0	293	6	Abu81097 Human PRO
21	1608	100.0	293	6	Abu72540 Novel hum
22	1608	100.0	293	6	Abu66797 Human PRO
23	1608	100.0	293	6	Ada41303 Human sec
24	1608	100.0	293	6	Ada41305 Human sec
25	1608	100.0	293	6	Ada40825 Human sec

26	1608	100.0	293	6	ADA41306 Human sec
27	1608	100.0	293	6	ABU59878 Novel sec
28	1608	100.0	293	6	ABU59289 Human sec
29	1608	100.0	293	6	ABO25986 Human PRO
30	1608	100.0	293	6	ABO25068 Human sec
31	1608	100.0	293	6	ABU58995 Human sec
32	1608	100.0	293	6	ABU92373 Novel hum
33	1608	100.0	293	6	ABU59438 Novel hum
34	1608	100.0	293	6	ABU67073 Human sec
35	1608	100.0	293	6	ABU92204 Novel hum
36	1608	100.0	293	6	ABU10910 Human PRO
37	1608	100.0	293	6	ABU81662 Novel hum
38	1608	100.0	293	6	ABU88601 Human sec
39	1608	100.0	293	6	ABO34115 Human PRO
40	1608	100.0	293	6	ADA45975 Novel hum
41	1608	100.0	293	6	ADA76406 Human PRO
42	1608	100.0	293	6	ADA19056 Human PRO
43	1608	100.0	293	6	ADA61679 Homo sapi
44	1608	100.0	293	6	ADB19464 Novel hum
45	1608	100.0	293	6	ADB28005 Human PRO

## ALIGNMENTS

### RESULT 1

RAY16777

ID AAY16777 standard; protein; 293 AA.

XX

AC AAY16777;

XX

DT 02-AUG-1999 (first entry)

XX

DE Human keratinocyte derived protease (KDP).

XX

KW Keratinocyte derived protease; KDP; protease; keratinocyte; human;  
KW skin care product; skin flaking; dandruff; laundry detergent composition;  
KW cleaning composition; dishwashing product.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Peptide

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

Novel human protease useful for treating or preventing skin flaking.

Claim 2; Page 32-34; 43pp; English.

This represents a human keratinocyte derived protease (KDP) polypeptide. The protease is substantially similar to a sequence encoded by the insert in plasmid PERM BP-6129. The KDP polypeptide, from human keratinocytes, is specifically used in skin care products, particularly to treat or prevent skin flaking, e.g. dandruff, and in laundry detergent compositions. More generally the KDP polypeptide can be used in any cleaning composition, e.g. hard surface or dishwashing products

```

XX SQ Sequence 293 AA;
Query Match 100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Qy 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPGRPVGVTNLCKFTKWIQETIQANS 293
Db 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPGRPVGVTNLCKFTKWIQETIQANS 293

RESULT 2
AAY30524
ID AAY30524 standard; protein; 293 AA.
AC AAY30524;
DT 03-DEC-1999 (first entry)
DE Human PDSP-1 protein.
KW PDSP-1; Tango 114; prostate-derived serine protease; fusion protein;
KW treatment; proliferative disorder; prostate cancer; screening;
KW predictive medicine; growth factor; growth factor biosynthesis;
KW cellular proliferation; growth factor binding protein; metastasis;
KW cellular differentiation; prostate development; detection; forensic;
KW serine protease; human.
XX Homo sapiens.
XX WO9946391-A2.
XX 16-SEP-1999.
XX 12-MAR-1999; 99WO-US005416.
XX 12-MAR-1998; 98US-00041400.
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX Holtzman DA;
XX WPI; 1999-561678/47.
XX N-PSDB; AAZ23317, AAZ23318.
XX Novel prostate derived serine protease polynucleotides and polypeptides
XX used to modulate cellular processes.
XX Claim 1c; Fig 1A; 94pp; English.
XX This invention describes a novel human prostate-derived serine protease
XX (PDSP) polynucleotide (also known as Tango 114). The prostate derived
XX serine protease (PDSP) polypeptide is used to modulate a variety of
XX cellular processes. It can be used to produce fusion proteins. PDSP is used
XX to treat proliferative disorders, e.g. prostate cancer. The protein may
XX also be used to produce antibodies, and to identify antagonists and

```

```

CC agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies
CC can be used in screening assays; predictive medicine; and methods of
CC treatment. PDSP cleaves growth factors, and can be used for the
CC modulation of growth factor biosynthesis; generation of active peptides;
CC regulation of cellular proliferation; degradation of growth factor
CC binding proteins; regulation of cellular differentiation; regulation of
CC metastasis; and regulation of prostate development. The PDSP
CC polynucleotides can be used to express the protein; to detect PDSP mRNA;
CC to detect genetic alterations in the PDSP gene; in forensic biology; and
CC as a source of primers and probes. As serine protease have important
CC roles in cellular processes, there exists a need for identifying novel
CC serine protease such as the prostate derived serine protease (PDSP) of
CC the invention. This sequence represents the human PDSP-1 described in the
CC invention
XX
XX SQ Sequence 293 AA;
Query Match 100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MATARPPMMVLCALITALLLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLLLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTQPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPS 180
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHFPKVLQCLNISVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240

Qy 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPGRPVGVTNLCKFTKWIQETIQANS 293
Db 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPGRPVGVTNLCKFTKWIQETIQANS 293

```

```

RESULT 3
AAY38412
ID AAY38412 standard; protein; 293 AA.
AC AAY38412;
DT 30-SEP-1999 (first entry)
DE Human secreted protein encoded by gene No. 27.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX Homo sapiens.
XX WO9935158-A1.
XX 15-JUL-1999.
XX 06-JAN-1999; 99WO-US000108.
XX 07-JAN-1998; 98US-0070657P.
XX 07-JAN-1998; 98US-0070658P.
XX 07-JAN-1998; 98US-0070692P.
XX 07-JAN-1998; 98US-0070704P.

```

```

XX (HUMA-) HUMAN GENOME SCI INC.
XX PA
XX Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;
XX PI Olsen HS, Duan RD, Rosen CA;
XX PD
XX WPI; 1999-444190/37.
XX DR N-PSDB; AA206245.
XX PT
XX New isolated human genes and the secreted polypeptides they encode.
XX PS
XX Claim 11; Page 188-189; 227pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AA206210) for increasing the stability of the fused protein
XX as compared to the human protein only. The invention relates to 36 novel
XX genes and their fragments (nucleic acid sequences: AA206219-206263; amino
XX acid sequences AAY38386-Y38498) which are useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX pathological conditions can be diagnosed by determining the amount of the
XX new polypeptides in a sample or by determining the presence of mutations
XX in the new polynucleotides. Specific uses are described for each of the
XX 36 polynucleotides, based on which tissues they are most highly expressed
XX in (see AA206219 for described uses)
XX
XX Sequence 293 AA;
XX
Query Match          100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSDCHPSNTVPVSGNODLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSDCHPSNTVPVSGNODLGAGAGEDARS 60
QY 61 DDSSRIINGSCDMDHTOPWQAALLRPNQLYCGAVLVHPQWLLTAACHRCRKKVFRVLGH 120
DB 61 DDSSRIINGSCDMDHTOPWQAALLRPNQLYCGAVLVHPQWLLTAACHRCRKKVFRVLGH 120
QY 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPS 180
DB 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPS 180
QY 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
DB 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
RESULT 4
AAY38426
ID AAY38426 standard; protein; 293 AA.
XX
XX AAY38426;
XX
XX 30-SEP-1999 (first entry)
XX
XX Human secreted protein encoded by gene No. 27.
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX developmental abnormality; foetal deficiency; blood; allergy; renal;
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

```

```

OS Homo sapiens.
XX
XX WO9935158-A1.
XX
XX 15-JUL-1999.
XX
XX 06-JAN-1999; 99WO-US000108.
XX
XX 07-JAN-1998; 98US-0070657P.
XX
XX 07-JAN-1998; 98US-0070658P.
XX
XX 07-JAN-1998; 98US-0070692P.
XX
XX 07-JAN-1998; 98US-0070704P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Soppet DR, Ebner R, Lafleur DW, Ni J, Brewer LA;
XX PI Olsen HS, Duan RD, Rosen CA;
XX
XX WPI; 1999-444190/37.
XX
XX N-PSDB; AA206259.
XX
XX New isolated human genes and the secreted polypeptides they encode.
XX
XX Claim 11; Page 200-201; 227pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AA206210) for increasing the stability of the fused protein
XX as compared to the human protein only. The invention relates to 36 novel
XX genes and their fragments (nucleic acid sequences: AA206219-206263; amino
XX acid sequences AAY38386-Y38498) which are useful for preventing, treating
XX or ameliorating medical conditions e.g. by protein or gene therapy. Also,
XX pathological conditions can be diagnosed by determining the amount of the
XX new polypeptides in a sample or by determining the presence of mutations
XX in the new polynucleotides. Specific uses are described for each of the
XX 36 polynucleotides, based on which tissues they are most highly expressed
XX in (see AA206219 for described uses)
XX
XX Sequence 293 AA;
XX
Query Match          100.0%; Score 1608; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSDCHPSNTVPVSGNODLGAGAGEDARS 60
DB 1 MATARPPMMVLCALITALLLGVTGTEHVLANNVSDCHPSNTVPVSGNODLGAGAGEDARS 60
QY 61 DDSSRIINGSCDMDHTOPWQAALLRPNQLYCGAVLVHPQWLLTAACHRCRKKVFRVLGH 120
DB 61 DDSSRIINGSCDMDHTOPWQAALLRPNQLYCGAVLVHPQWLLTAACHRCRKKVFRVLGH 120
QY 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPS 180
DB 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHNDLMLIKLNRRIPTKDVPRINVSCHPS 180
QY 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGWGTTSKPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
DB 241 CQDSDGGPVVNCVSGLSQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
RESULT 5
AAY66726
ID AAY66726 standard; protein; 293 AA.
XX
XX AAY66726;
XX
XX 05-APR-2000 (first entry)

```

XX	Membrane-bound protein PRO1132.	PR	24-JUN-1998;	98US-0090461P.
DE		PR	24-JUN-1998;	98US-0090472P.
XX		PR	24-JUN-1998;	98US-0090535P.
XX	Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;	PR	24-JUN-1998;	98US-0090538P.
KW	pharmaceutical; receptor immunoadhesin; gene mapping.	PR	24-JUN-1998;	98US-0090540P.
KW		PR	24-JUN-1998;	98US-0090557P.
XX		PR	25-JUN-1998;	98US-0090676P.
OS	Homo sapiens.	PR	25-JUN-1998;	98US-0090678P.
XX	WO9963088-A2.	PR	25-JUN-1998;	98US-0090688P.
PN		PR	25-JUN-1998;	98US-0090690P.
XX		PR	25-JUN-1998;	98US-0090691P.
PD	09-DEC-1999.	PR	25-JUN-1998;	98US-0090694P.
XX		PR	25-JUN-1998;	98US-0090695P.
PF	02-JUN-1999;	PR	25-JUN-1998;	98US-0090696P.
XX		PR	26-JUN-1998;	98US-0090862P.
XX	02-JUN-1998;	PR	26-JUN-1998;	98US-0090863P.
PR	98US-0087607P.	PR	01-JUL-1998;	98US-0091358P.
PR	98US-0087609P.	PR	01-JUL-1998;	98US-0091360P.
PR	98US-0087759P.	PR	02-JUL-1998;	98US-0091478P.
PR	98US-0087827P.	PR	02-JUL-1998;	98US-0091486P.
PR	98US-0088021P.	PR	02-JUL-1998;	98US-0091519P.
PR	98US-0088025P.	PR	02-JUL-1998;	98US-0091544P.
PR	98US-0088028P.	PR	02-JUL-1998;	98US-0091626P.
PR	98US-0088029P.	PR	02-JUL-1998;	98US-0091628P.
PR	98US-0088030P.	PR	02-JUL-1998;	98US-0091633P.
PR	98US-0088033P.	PR	02-JUL-1998;	98US-0091646P.
PR	98US-0088326P.	PR	02-JUL-1998;	98US-0091673P.
PR	98US-0088167P.	PR	02-JUL-1998;	98US-0091788P.
PR	98US-0088202P.	PR	07-JUL-1998;	98US-0091982P.
PR	98US-0088212P.	PR	07-JUL-1998;	98US-0092182P.
PR	98US-0088217P.	PR	09-JUL-1998;	98US-0092472P.
PR	98US-0088555P.	PR	10-JUL-1998;	98US-0093339P.
PR	98US-0088722P.	PR	20-JUL-1998;	98US-0094651P.
PR	98US-0088730P.	PR	30-JUL-1998;	98US-0095282P.
PR	98US-0088734P.	PR	04-AUG-1998;	98US-0095285P.
PR	98US-0088738P.	PR	04-AUG-1998;	98US-0095301P.
PR	98US-0088740P.	PR	04-AUG-1998;	98US-0095302P.
PR	98US-0088741P.	PR	04-AUG-1998;	98US-0095318P.
PR	98US-0088742P.	PR	04-AUG-1998;	98US-0095321P.
PR	98US-0088810P.	PR	04-AUG-1998;	98US-0095325P.
PR	98US-0088811P.	PR	10-AUG-1998;	98US-0095916P.
PR	98US-0088824P.	PR	10-AUG-1998;	98US-0095929P.
PR	98US-0088825P.	PR	10-AUG-1998;	98US-0096012P.
PR	98US-0088826P.	PR	11-AUG-1998;	98US-0096143P.
PR	98US-0088858P.	PR	11-AUG-1998;	98US-0096146P.
PR	98US-0088861P.	PR	12-AUG-1998;	98US-0096329P.
PR	98US-0088863P.	PR	17-AUG-1998;	98US-0096757P.
PR	98US-0088876P.	PR	17-AUG-1998;	98US-0096766P.
PR	98US-0089090P.	PR	17-AUG-1998;	98US-0096768P.
PR	98US-0089105P.	PR	17-AUG-1998;	98US-0096773P.
PR	98US-0089440P.	PR	17-AUG-1998;	98US-0096791P.
PR	98US-0089512P.	PR	17-AUG-1998;	98US-0096867P.
PR	98US-0089513P.	PR	17-AUG-1998;	98US-0096891P.
PR	98US-0089532P.	PR	17-AUG-1998;	98US-0096894P.
PR	98US-0089538P.	PR	17-AUG-1998;	98US-0096895P.
PR	98US-0089598P.	PR	17-AUG-1998;	98US-0096897P.
PR	98US-0089599P.	PR	18-AUG-1998;	98US-0096949P.
PR	98US-0089600P.	PR	18-AUG-1998;	98US-0096950P.
PR	98US-0089653P.	PR	18-AUG-1998;	98US-0096959P.
PR	98US-0089801P.	PR	18-AUG-1998;	98US-0096960P.
PR	98US-0089907P.	PR	18-AUG-1998;	98US-0097022P.
PR	98US-0089908P.	PR	18-AUG-1998;	98US-0097141P.
PR	98US-0089947P.	PR	19-AUG-1998;	98US-0097218P.
PR	98US-0089948P.	PR	20-AUG-1998;	98US-0097661P.
PR	98US-0089952P.	PR	24-AUG-1998;	98US-0097661P.
PR	98US-0090248P.	PR	26-AUG-1998;	98US-0097951P.
PR	98US-0090252P.	PR	26-AUG-1998;	98US-0097952P.
PR	98US-0090254P.	PR	26-AUG-1998;	98US-0097954P.
PR	98US-0090349P.	PR	26-AUG-1998;	98US-0097955P.
PR	98US-0090355P.	PR	26-AUG-1998;	98US-0097971P.
PR	98US-0090429P.	PR	26-AUG-1998;	98US-0097974P.
PR	98US-0090431P.	PR	26-AUG-1998;	98US-0097978P.
PR	98US-0090433P.	PR	26-AUG-1998;	98US-0097979P.
PR	98US-0090444P.	PR	26-AUG-1998;	98US-0097986P.
PR	98US-0090445P.			

PR 26-AUG-1998; 98US-0098014P.  
 PR 31-AUG-1998; 98US-0098525P.  
 PR 16-SEP-1998; 98US-0100634P.  
 PR 12-JAN-1999; 99US-0115565P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 PA  
 XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
 PI Wood WI, Yuan J;  
 XX  
 XX WPI; 2000-072883/06.  
 DR N-PSDB; AA265070.  
 XX  
 XX Membrane-bound proteins and related nucleotide sequences.  
 XX  
 XX Claim 12; Fig 226; 822pp; English.  
 PS  
 XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptors, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences will  
 CC also be useful for the preparation of PRO polypeptides, especially by  
 CC recombinant techniques  
 XX  
 XX Sequence 293 AA;  
 SQ  
 Query Match 100.0%; Score 1608; DB 3; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60  
 DB 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60  
 QY 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAAHCKKVFVRVLGH 120  
 DB 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAAHCKKVFVRVLGH 120  
 QY 121 YLSLSPVYESGGQMFQGVKSIHPHGYSHPGHNDMLIKLNRRIRPTKDVPRPINVSSHCP 180  
 DB 121 YLSLSPVYESGGQMFQGVKSIHPHGYSHPGHNDMLIKLNRRIRPTKDVPRPINVSSHCP 180  
 QY 181 AGTKCLVSGWGTTSKSPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 DB 181 AGTKCLVSGWGTTSKSPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 QY 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARNRPGVVTNLCKFTKTIQETIQANS 293  
 DB 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARNRPGVVTNLCKFTKTIQETIQANS 293  
 RESULT 6  
 AAB21296  
 ID AAB21296 standard; protein; 293 AA.  
 XX  
 XX AAB21296;  
 AC  
 XX 02-FEB-2001 (first entry)  
 DT  
 XX Human KLK-L2 protein.  
 DE  
 XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
 KW kallikrein-like protein; serine protease; cytostatic; cancer;  
 KW prostrate cancer.  
 XX

OS Homo sapiens.  
 XX WO200053776-A2.  
 PN  
 XX 14-SEP-2000.  
 PD  
 XX 09-MAR-2000; 2000WO-CA000258.  
 PF  
 XX 11-MAR-1999; 99US-0124260P.  
 PR 01-APR-1999; 99US-0127386P.  
 PR 21-JUL-1999; 99US-0144919P.  
 XX  
 XX (MOUN ) MOUNT SINAI HOSPITAL.  
 PA  
 XX Yousef GM, Diamandis EP;  
 PI  
 XX WPI; 2000-587440/55.  
 DR N-PSDB; AAA95905.  
 XX  
 XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
 PT protein mediated disorders, especially cancer.  
 XX  
 XX Claim 9; Page 149-150; 184pp; English.  
 PS  
 XX The present sequence is kallikrein-like protein KLK-L2. Kallikreins and  
 CC kallikrein-like proteins are a subgroup of the serine protease enzyme  
 CC family. They catalyze the selective cleavage of specific polypeptide  
 CC precursors to release peptides with potent biological activity. Nucleic  
 CC acids encoding kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4,  
 CC KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the  
 CC treatment, monitoring and diagnosis of cancers, especially prostate  
 CC cancer. They can also be used to identify a substance that can associate  
 CC with or mediate the biological activity of the proteins. Antibodies can  
 CC be used to treat conditions mediated by the kallikrein-like proteins  
 XX  
 XX Sequence 293 AA;  
 SQ  
 Query Match 100.0%; Score 1608; DB 3; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60  
 DB 1 MATARPPMMWVLCALITALLIGVTEHVLANNVSCDHPSTNTVPVSGNODLGAGAGEDARS 60  
 QY 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAAHCKKVFVRVLGH 120  
 DB 61 DSSSRIINGSDCMHTOPWQAALLRLPNOLYCGAVLVHPQWLLTAAHCKKVFVRVLGH 120  
 QY 121 YLSLSPVYESGGQMFQGVKSIHPHGYSHPGHNDMLIKLNRRIRPTKDVPRPINVSSHCP 180  
 DB 121 YLSLSPVYESGGQMFQGVKSIHPHGYSHPGHNDMLIKLNRRIRPTKDVPRPINVSSHCP 180  
 QY 181 AGTKCLVSGWGTTSKSPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 DB 181 AGTKCLVSGWGTTSKSPQVHFPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 QY 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARNRPGVVTNLCKFTKTIQETIQANS 293  
 DB 241 CQDGGGPPVNCVNGSLQGLVSWGDYPCARNRPGVVTNLCKFTKTIQETIQANS 293  
 RESULT 7  
 AAU12399  
 ID AAU12399 standard; protein; 293 AA.  
 XX  
 XX AAU12399;  
 AC  
 XX 24-OCT-2001 (first entry)  
 DT  
 XX Human PRO1132 polypeptide sequence.  
 DE  
 XX Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast;  
 KW

KW prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage;  
KW ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte;  
KW A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.

XX WO200140466-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US032678.

XX 01-DEC-1999; 99WO-US028301.

XX 01-DEC-1999; 99WO-US028634.

XX 02-DEC-1999; 99WO-US028551.

XX 02-DEC-1999; 99WO-US028564.

XX 02-DEC-1999; 99WO-US028565.

XX 09-DEC-1999; 99US-0170262P.

XX 16-DEC-1999; 99WO-US030095.

XX 20-DEC-1999; 99WO-US030911.

XX 20-DEC-1999; 99WO-US030999.

XX 30-DEC-1999; 99WO-US031243.

XX 30-DEC-1999; 99WO-US031274.

XX 05-JAN-2000; 2000WO-US000219.

XX 06-JAN-2000; 2000WO-US000277.

XX 11-FEB-2000; 2000WO-US000376.

XX 18-FEB-2000; 2000WO-US003565.

XX 18-FEB-2000; 2000WO-US004341.

XX 22-FEB-2000; 2000WO-US004342.

XX 22-FEB-2000; 2000WO-US004414.

XX 24-FEB-2000; 2000WO-US004914.

XX 01-MAR-2000; 2000WO-US005004.

XX 02-MAR-2000; 2000WO-US005601.

XX 03-MAR-2000; 2000WO-US005841.

XX 10-MAR-2000; 2000US-0187202P.

XX 15-MAR-2000; 2000WO-US006319.

XX 20-MAR-2000; 2000WO-US006884.

XX 21-MAR-2000; 2000WO-US007377.

XX 30-MAR-2000; 2000WO-US007532.

XX 17-MAY-2000; 2000WO-US008439.

XX 22-MAY-2000; 2000WO-US013705.

XX 30-MAY-2000; 2000WO-US014042.

XX 02-JUN-2000; 2000WO-US014941.

XX 05-JUN-2000; 2000WO-US015264.

XX 28-JUL-2000; 2000US-0209832P.

CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample. Some  
CC of the 275 sequences are also useful to stimulate the release of tumour  
CC necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or  
CC differentiation of chondrocytes, the proliferation or gene expression in  
CC pericyte cells, the release of proteoglycans from cartilage, the  
CC proliferation of inner ear utricular supporting cells or of T-  
CC lymphocytes, the release of a cytokine from peripheral blood monocytes  
CC (PBMCs), or the proliferation of endothelial cells. Some of the PRO  
CC polypeptides may modulate glucose or free fatty acid uptake by skeletal  
CC muscle cells or by adipocytes; or inhibit binding of A-peptide to factor  
CC VIIA. The PRO polypeptides can be used in assays to identify molecules  
CC involved in binding interactions. The polynucleotides encoding PRO  
CC polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy  
XX  
SQ Sequence 293 AA;

Query Match 100.0%; Score 1608; DB 4; Length 293;  
Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTVPSGSNODLGAGAGEDARS 60  
Db 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTVPSGSNODLGAGAGEDARS 60  
Qy 61 DSSSRIINGSDDMHTQPMQALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
Db 61 DSSSRIINGSDDMHTQPMQALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
Qy 121 YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDMLIKLNRIRPTKDVRIINVSCHCPS 180  
Db 121 YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDMLIKLNRIRPTKDVRIINVSCHCPS 180  
Qy 181 AGTKCLVSGWTTKSPQVHFVKVLCQLNIVLSOKRCEDAYPROIDDTMFCAKGRDS 240  
Db 181 AGTKCLVSGWTTKSPQVHFVKVLCQLNIVLSOKRCEDAYPROIDDTMFCAKGRDS 240  
Qy 241 CQDGGPVPVCGNSLQGLVSWGDYPCARPNRPGVYTNLCRFTKWIQETIQANS 293  
Db 241 CQDGGPVPVCGNSLQGLVSWGDYPCARPNRPGVYTNLCRFTKWIQETIQANS 293

RESULT 8

AAB65249

ID AAB65249 standard; protein; 293 AA.

XX AAB65249;

AC AAB65249;

XX 02-APR-2001 (first entry)

XX Human PRO1132 (UNQ570) protein sequence SEQ ID NO:309.

DE Human; secreted and transmembrane protein; PRO; cytostatic; cell death;  
KW cancer; chromosomal mapping; gene mapping; tissue typing;  
KW diagnostic assay.

XX Homo sapiens.

OS WO200073454-A1.

XX 07-DEC-2000.

XX 30-MAR-2000; 2000WO-US008439.

XX 02-JUN-1999; 99WO-US012252.

XX 23-JUN-1999; 99US-0141037P.

XX 07-JUL-1999; 99US-0143048P.

XX 20-JUL-1999; 99US-0144758P.

XX 26-JUL-1999; 99US-0145698P.

XX 28-JUL-1999; 99US-0146222P.

XX 17-AUG-1999; 99US-0149396P.

XX 15-SEP-1999; 99WO-US021090.



PR 15-SEP-1999; 99WO-US021547.  
 PR 08-OCT-1999; 99US-0158663P.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 16-DEC-1999; 99WO-US030095.  
 PR 20-DEC-1999; 99WO-US030911.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US000356.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US004914.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 20-MAR-2000; 2000WO-US007377.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
 PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
 PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI,  
 PI Zhang Z;  
 XX  
 WPI: 2001-032160/04.  
 DR N-PSDB; AAF44216.  
 XX  
 PRO polynucleotides used to produce polypeptides used to target bioactive  
 molecules such as toxins, radiolabels or antibodies, to specific cells,  
 to cause targeted cell death.  
 PT  
 PT  
 PT  
 PT  
 XX  
 XX  
 PS Claim 12; Fig 226; 935pp; English.  
 XX  
 The present invention describes human secreted and transmembrane PRO  
 proteins. The PRO proteins have cytostatic activity. The PRO proteins can  
 be used for targeted delivery of bioactive molecules, such as toxins,  
 radiolabels or antibodies, that cause cell death. PRO nucleotide  
 sequences, and their fragments, can be used as hybridisation probes, in  
 chromosomal and gene mapping, and in the generation of anti-sense RNA and  
 DNA. They may also be used to produce transgenic animals which are used  
 to develop and screen therapeutically useful reagents. The PRO nucleotide  
 and protein sequence can be used for tissue typing and in treating  
 cancer. Anti-PRO antibodies can be used in diagnostic assays. AAF44270 to  
 AAF44470 represent PCR primers and hybridisation probes used in the  
 isolation of human PRO sequences. AAF44087 to AAF44269 and AAF65154 to  
 AAF65300 represent human PRO polynucleotide and protein sequences given  
 in the exemplification of the present invention  
 CC  
 XX Sequence 293 AA;  
 SQ  
 Query Match 100.0%; Score 1608; DB 4; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATARPPMMVLCALITALLGVTEHVLANDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
 DB 1 MATARPPMMVLCALITALLGVTEHVLANDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
 QY 61 DSSSRITNGSDCMHTQWQAALLRPNOLYCGAVLHPQWLLTAAHCKKVFVRVLGH 120  
 DB 61 DSSSRITNGSDCMHTQWQAALLRPNOLYCGAVLHPQWLLTAAHCKKVFVRVLGH 120  
 QY 121 YLSLSPVYESGQMFQGVKSIPHPGYSHPGSHNDMLIKLNRIRTPKDVPRINVSCHPS 180  
 DB 121 YLSLSPVYESGQMFQGVKSIPHPGYSHPGSHNDMLIKLNRIRTPKDVPRINVSCHPS 180  
 QY 181 AGTKCLVSGWGTTSQPQVHPFKVLQCLNISVLISQKRCEDAYPRQIDDTMFCAKGRDSD 240  
 DB 181 AGTKCLVSGWGTTSQPQVHPFKVLQCLNISVLISQKRCEDAYPRQIDDTMFCAKGRDSD 240  
 QY 241 CQDGGGPPVVCNGSLQGLVSMGDDYPCARPNGPVYTNLCCKTKWIQETIQANS 293  
 DB 241 CQDGGGPPVVCNGSLQGLVSMGDDYPCARPNGPVYTNLCCKTKWIQETIQANS 293

---

Db 241 CQDGGGPPVVCNGSLQGLVSMGDDYPCARPNGPVYTNLCCKTKWIQETIQANS 293  
 RESULT 9  
 AAU81966  
 ID AAU81966 standard; protein; 293 AA.  
 XX  
 AC AAU81966;  
 XX  
 DT 09-APR-2002 (first entry)  
 DE Human PRO1132.  
 XX  
 KW Human; PRO: antiinflammatory; ophthalmological; vasotropic;  
 KW retinal cell injury; ocular disease; retinitis pigmentosa;  
 KW macular degeneration; retinal detachment; retinal tear; retinopathy;  
 KW retinal degenerative disease; macular hole; degenerative myopia;  
 KW acute retinal necrosis syndrome; traumatic chorioretinopathy;  
 KW Purtscher's retinopathy; oedema; ischaemic condition;  
 KW retinal vision occlusion; collagen vascular disease;  
 KW thrombocytopenic purpura; uveitis; retinal vasculitis; Eales disease;  
 KW systemic lupus erythematosus; environmental trauma.  
 OS  
 XX Homo sapiens.  
 XX WO200109327-A2.  
 PN  
 XX  
 PD 08-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000WO-US020710.  
 XX  
 PR 28-JUL-1999; 99US-0146222P.  
 PR 13-SEP-1999; 99WO-US020944.  
 PR 15-SEP-1999; 99WO-US021090.  
 PR 29-NOV-1999; 99WO-US028214.  
 PR 30-NOV-1999; 99WO-US028313.  
 PR 01-DEC-1999; 99WO-US028301.  
 PR 05-JAN-2000; 2000WO-US000219.  
 PR 06-JAN-2000; 2000WO-US000376.  
 PR 11-FEB-2000; 2000WO-US003565.  
 PR 18-FEB-2000; 2000WO-US004341.  
 PR 22-FEB-2000; 2000WO-US004414.  
 PR 24-FEB-2000; 2000WO-US005004.  
 PR 02-MAR-2000; 2000WO-US005841.  
 PR 15-MAR-2000; 2000WO-US006884.  
 PR 30-MAR-2000; 2000WO-US008439.  
 PR 17-MAY-2000; 2000WO-US013705.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Baker KP, Goddard A, Godowski PJ, Gurney AL;  
 PI Kljavin IJ, Lafleur M, Mark MR, Marsters SA, Pitti RM, Watanabe CK;  
 PI Wood WI;  
 XX  
 DR WPI: 2002-130120/17.  
 DR N-PSDB; ABK28605.  
 XX  
 PT Promoting survival of retinal cells, or delaying or preventing retinal  
 PT cell injury or death, by contacting retinal cells with PRO175, 220, 216,  
 PT 243, 306, 346, 322, 536, 943, 840, 828, 826, 1068 or PRO1132 polypeptide.  
 XX  
 PS Claim 44; Fig 29; 152pp; English.  
 XX  
 The invention relates to promoting the survival of retinal cells, or  
 CC delaying or preventing retinal cell injury or death, by contacting the  
 CC retinal cells with the polypeptide such as PRO175, PRO220, PRO216,  
 CC PRO243, PRO306, PRO346, PRO322, PRO336, PRO943, PRO840, PRO828, PRO826,  
 CC PRO1068 or PRO1132 polypeptide. Also included are the nucleic acids  
 CC encoding the PRO proteins, a vector comprising the nucleic acid, a host  
 CC cell comprising the vector, and anti-PRO antibody. The PRO proteins are  
 CC useful for promoting survival of retinal cells (retinal neurons such as  
 CC retinal ganglion cells, displaced retinal ganglion cells, amacrine cells,  
 CC displaced amacrine cells, horizontal neurons or bipolar neurons, rod

```
CC photoreceptors, or supportive cells such as Muller cells or pigment
CC epithelial cells), or delaying or preventing retinal cell injury or death
CC caused by ocular disease (which is or is associated with retinitis
CC pigmentosa, macular degeneration, retinal detachment, retinal tear,
CC retinopathy, retinal degenerative disease, macular hole, degenerative
CC myopia, acute retinal necrosis syndrome, traumatic chorioretinopathy or
CC contusion, Purtscher's retinopathy, oedema, an ischaemic condition,
CC central or branch retinal vision occlusion, collagen vascular disease,
CC thrombocytopenic purpura, uveitis, retinal vasculitis, occlusion
CC associated with Bales disease or systemic lupus erythematosus), retinal
CC injury or environmental trauma. The retinal cell injury or death is
CC delayed or prevented by substantially not causing angiogenesis or
CC mitogenesis. The present sequence represents a PRO protein
XX
SQ Sequence 293 AA;

Query Match 100.0%; Score 1608; DB 5; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Dy 1 MATARPPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRIINGSDCMHTOPWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Dy 61 DSSSRIINGSDCMHTOPWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMGQGVKSIHPHGYSHGHNDMLKLNRIIRPTKDVRIINVSCHPS 180
Dy 121 YSLSPVYESGQMGQGVKSIHPHGYSHGHNDMLKLNRIIRPTKDVRIINVSCHPS 180

Qy 181 AGTKCLVSGWGTTSQVHFPPKVLQCLNLSVLQKRCEDAYPROIDDTMFCAKGRDSDS 240
Dy 181 AGTKCLVSGWGTTSQVHFPPKVLQCLNLSVLQKRCEDAYPROIDDTMFCAKGRDSDS 240

Qy 241 CQSGSGGPPVNCVSLQGLVSGDYPFCARPNNPGVVTNLCKFTKWIQETIQANS 293
Dy 241 CQSGSGGPPVNCVSLQGLVSGDYPFCARPNNPGVVTNLCKFTKWIQETIQANS 293

RESULT 10
ID ABU58064
XX ABU58064 standard; protein; 293 AA.
AC ABU58064;
XX
DT 14-APR-2003 (first entry)
XX
DE Human PRO polypeptide #96.
XX
KW Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach; liver;
KW horse; cow; dog; cat; sheep; pig; goat; rabbit; ADEPT;
KW antibody-dependent enzyme mediated prodrug therapy.
XX
OS Homo sapiens.
XX
PN US2003027163-A1.
XX
PD 06-FEB-2003.
XX
PF 15-NOV-2001; 2001US-00997666.
XX
PR 16-JUN-1997; 97US-0049787P.
PR 17-OCT-1997; 97US-0062250P.
PR 05-NOV-1997; 97WO-US020069.
PR 12-NOV-1997; 97US-0065186P.
PR 13-NOV-1997; 97US-0065311P.
PR 24-NOV-1997; 97US-0066770P.
PR 25-FEB-1998; 98US-0075945P.
PR 20-MAR-1998; 98US-0078910P.
PR 28-APR-1998; 98US-0083322P.
PR 07-MAY-1998; 98US-0084600P.

28-MAY-1998; 98US-0087106P.
02-JUN-1998; 98US-0087607P.
02-JUN-1998; 98US-0087609P.
02-JUN-1998; 98US-0087759P.
03-JUN-1998; 98US-0087827P.
04-JUN-1998; 98US-0088021P.
04-JUN-1998; 98US-0088025P.
04-JUN-1998; 98US-0088026P.
04-JUN-1998; 98US-0088028P.
04-JUN-1998; 98US-0088029P.
04-JUN-1998; 98US-0088030P.
04-JUN-1998; 98US-0088033P.
04-JUN-1998; 98US-0088326P.
05-JUN-1998; 98US-0088167P.
05-JUN-1998; 98US-0088202P.
05-JUN-1998; 98US-0088212P.
05-JUN-1998; 98US-0088217P.
09-JUN-1998; 98US-0088655P.
10-JUN-1998; 98US-0088734P.
10-JUN-1998; 98US-0088738P.
10-JUN-1998; 98US-0088742P.
10-JUN-1998; 98US-0088810P.
10-JUN-1998; 98US-0088824P.
10-JUN-1998; 98US-0088826P.
11-JUN-1998; 98US-0088858P.
11-JUN-1998; 98US-0088861P.
11-JUN-1998; 98US-0088876P.
12-JUN-1998; 98US-0089105P.
16-JUN-1998; 98US-0089440P.
16-JUN-1998; 98US-0089512P.
16-JUN-1998; 98US-0089514P.
17-JUN-1998; 98US-0089532P.
17-JUN-1998; 98US-0089538P.
17-JUN-1998; 98US-0089598P.
17-JUN-1998; 98US-0089599P.
17-JUN-1998; 98US-0089600P.
17-JUN-1998; 98US-0089653P.
18-JUN-1998; 98US-0089801P.
18-JUN-1998; 98US-0089907P.
18-JUN-1998; 98US-0089908P.
19-JUN-1998; 98US-0089947P.
19-JUN-1998; 98US-0089948P.
19-JUN-1998; 98US-0089952P.
22-JUN-1998; 98US-0090246P.
22-JUN-1998; 98US-0090252P.
23-JUN-1998; 98US-0090254P.
23-JUN-1998; 98US-0090349P.
23-JUN-1998; 98US-0090355P.
24-JUN-1998; 98US-0090429P.
24-JUN-1998; 98US-0090431P.
24-JUN-1998; 98US-0090435P.
24-JUN-1998; 98US-0090444P.
24-JUN-1998; 98US-0090445P.
24-JUN-1998; 98US-0090472P.
24-JUN-1998; 98US-0090535P.
24-JUN-1998; 98US-0090540P.
24-JUN-1998; 98US-0090542P.
24-JUN-1998; 98US-0090557P.
25-JUN-1998; 98US-0090676P.
25-JUN-1998; 98US-0090678P.
25-JUN-1998; 98US-0090690P.
25-JUN-1998; 98US-0090694P.
25-JUN-1998; 98US-0090695P.
26-JUN-1998; 98US-0090696P.
26-JUN-1998; 98US-0090862P.
26-JUN-1998; 98US-0090863P.
01-JUL-1998; 98US-0091360P.
01-JUL-1998; 98US-0091544P.
02-JUL-1998; 98US-0091478P.
02-JUL-1998; 98US-0091519P.
02-JUL-1998; 98US-0091626P.
02-JUL-1998; 98US-0091628P.
02-JUL-1998; 98US-0091633P.
```



PN US2002132252-A1.  
XX 19-SEP-2002.  
PD 14-NOV-2001; 2001US-00990442.  
XX 16-JUN-1997; 97US-0049787P.  
PR 17-OCT-1997; 97US-0062257P.  
PR 05-NOV-1997; 97WO-US020069.  
PR 12-NOV-1997; 97US-0065186P.  
PR 13-NOV-1997; 97US-0085311P.  
PR 24-NOV-1997; 97US-0066770P.  
PR 25-FEB-1998; 98US-0075945P.  
PR 20-MAR-1998; 98US-0078919P.  
PR 28-APR-1998; 98US-0083322P.  
PR 07-MAY-1998; 98US-0084600P.  
PR 28-MAY-1998; 98US-0087106P.  
PR 02-JUN-1998; 98US-0087607P.  
PR 02-JUN-1998; 98US-0087609P.  
PR 02-JUN-1998; 98US-0087759P.  
PR 03-JUN-1998; 98US-0087827P.  
PR 04-JUN-1998; 98US-0088021P.  
PR 04-JUN-1998; 98US-0088025P.  
PR 04-JUN-1998; 98US-0088026P.  
PR 04-JUN-1998; 98US-0088028P.  
PR 04-JUN-1998; 98US-0088029P.  
PR 04-JUN-1998; 98US-0088030P.  
PR 04-JUN-1998; 98US-0088033P.  
PR 04-JUN-1998; 98US-0088326P.  
PR 05-JUN-1998; 98US-0088167P.  
PR 05-JUN-1998; 98US-0088202P.  
PR 05-JUN-1998; 98US-0088212P.  
PR 05-JUN-1998; 98US-0088217P.  
PR 09-JUN-1998; 98US-0088655P.  
PR 10-JUN-1998; 98US-0088734P.  
PR 10-JUN-1998; 98US-0088738P.  
PR 10-JUN-1998; 98US-0088742P.  
PR 10-JUN-1998; 98US-0088810P.  
PR 10-JUN-1998; 98US-0088824P.  
PR 10-JUN-1998; 98US-0088826P.  
PR 11-JUN-1998; 98US-0088858P.  
PR 11-JUN-1998; 98US-0088861P.  
PR 11-JUN-1998; 98US-0088876P.  
PR 12-JUN-1998; 98US-0089105P.  
PR 16-JUN-1998; 98US-0089402P.  
PR 16-JUN-1998; 98US-0089512P.  
PR 16-JUN-1998; 98US-0089514P.  
PR 17-JUN-1998; 98US-0089532P.  
PR 17-JUN-1998; 98US-0089538P.  
PR 17-JUN-1998; 98US-0089598P.  
PR 17-JUN-1998; 98US-0089599P.  
PR 17-JUN-1998; 98US-0089600P.  
PR 17-JUN-1998; 98US-0089653P.  
PR 18-JUN-1998; 98US-0089801P.  
PR 18-JUN-1998; 98US-0089907P.  
PR 18-JUN-1998; 98US-0089908P.  
PR 16-SEP-1998; 98WO-US019330.  
PR 17-SEP-1998; 98WO-US019437.  
PR 07-OCT-1998; 98WO-US021141.  
PR 01-DEC-1998; 98WO-US025108.  
PR 05-JAN-1999; 98WO-US000106.  
PR 08-MAR-1999; 99WO-US005028.  
PR 02-JUN-1999; 99WO-US012252.  
PR 15-SEP-1999; 99WO-US021090.  
PR 15-SEP-1999; 99WO-US021547.  
PR 30-NOV-1999; 99WO-US028313.  
PR 01-DEC-1999; 99WO-US028301.  
PR 01-DEC-1999; 99WO-US028634.  
PR 16-DEC-1999; 99WO-US030095.  
PR 20-DEC-1999; 99WO-US030911.  
PR 06-JAN-2000; 2000WO-US000219.  
PR 06-JAN-2000; 2000WO-US000376.  
PR 11-FEB-2000; 2000WO-US003565.  
PR 18-FEB-2000; 2000WO-US004341.  
PR 22-FEB-2000; 2000WO-US004414.  
PR 24-FEB-2000; 2000WO-US004914.  
PR 24-FEB-2000; 2000WO-US005004.  
PR 02-MAR-2000; 2000WO-US005841.  
PR 10-MAR-2000; 2000WO-US006319.  
PR 15-MAR-2000; 2000WO-US006884.  
PR 20-MAR-2000; 2000WO-US007377.  
PR 30-MAR-2000; 2000WO-US008439.  
PR 15-MAY-2000; 2000WO-US013358.  
PR 17-MAY-2000; 2000WO-US013705.  
PR 22-MAY-2000; 2000WO-US014042.  
PR 30-MAY-2000; 2000WO-US014941.  
PR 02-JUN-2000; 2000WO-US015264.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 11-AUG-2000; 2000WO-US022031.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 08-NOV-2000; 2000WO-US030952.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 20-JUN-2001; 2001WO-US019692.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-JUL-2001; 2001WO-US021735.  
PR 28-AUG-2001; 2001US-00941992.  
XX (GETH ) GENENTECH INC.  
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NP;  
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams EM, Wood WT;  
PI Zhang Z;  
XX WPI; 2003-2477083/24.  
DR N-PSDB; ABX80317.  
XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
PT are therapeutically useful for enhancing immune response and in cancer  
PT treatments.  
XX Claim 12; Fig 226; 648pp; English.  
XX The invention describes an isolated human PRO polypeptide. The PRO  
CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
CC in modulating at least one biological activity of a cell expressing a PRO  
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
CC stimulate adrenal cortical capillary endothelial growth and PRO536,  
CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
CC useful for treating conditions or disorders where angiogenesis would be  
CC beneficial, e.g. wound healing and antagonist of this polypeptide are  
CC useful for treating cancerous tumours. PRO812 inhibits vascular  
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
CC cells and is thus useful for inhibiting endothelial cell growth in  
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC rod photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosum, AMD. PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or Crohn's  
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and are  
CC thus useful for treating sports injuries, and arthritis. This is the

CC amino acid sequence of a novel human PRO protein

XX SQ Sequence 293 AA;

Query Match 100.0%; Score 1608; DB 6; Length 293;  
Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMWLICALITALLIGVTEHVLANNVSCDHPSTNTVPSCNODLGAGAGEDARS 60  
DB 1 MATARPPMMWLICALITALLIGVTEHVLANNVSCDHPSTNTVPSCNODLGAGAGEDARS 60  
QY 61 DSSSRRIINGSDCMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
DB 61 DSSSRRIINGSDCMHTOPWQAALLLRPNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
QY 121 YSLSPVYESGQMFQGVKSIPHPGVSHPGSHNDMLIKLNRIRPTKQVRPINVSSHCP 180  
DB 121 YSLSPVYESGQMFQGVKSIPHPGVSHPGSHNDMLIKLNRIRPTKQVRPINVSSHCP 180  
QY 181 AGTKLVSGWGTTSQPQVHPKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
DB 181 AGTKLVSGWGTTSQPQVHPKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
QY 241 CQDSGGPVVNCNGSLQGLVSWGDPYPCARPNGPVYTNLCCKFTKWIQETIQANS 293  
DB 241 CQDSGGPVVNCNGSLQGLVSWGDPYPCARPNGPVYTNLCCKFTKWIQETIQANS 293

RESULT 12

ABU82654

ID ABU82654 standard; protein; 293 AA.

XX AC ABU82654;

XX DT 26-JUN-2003 (first entry)

XX DE Human secreted/transmembrane protein PRO1132.

XX KW Human; PRO; secreted protein; transmembrane protein;  
cardiac insufficiency disorders; angiogenesis; wound healing;  
cancerous tumour; immune response; retinal disorder; sight loss;  
retinitis pigmentosa; age-related macular degeneration; AMD;  
kidney disorder; Berger disease; nephropathy; dermatitis; herpetiformis;  
Crohn's disease; sports injury; arthritis.

XX OS Homo sapiens.

XX PN US2003032023-A1.

XX PD 13-FEB-2003.

XX PF 14-NOV-2001; 2001US-00990711.

XX PR 16-JUN-1997; 97US-0049787P.

PR 17-OCT-1997; 97US-0062250P.

PR 05-NOV-1997; 97WO-US020069.

PR 12-NOV-1997; 97US-0065186P.

PR 13-NOV-1997; 97US-0065311P.

PR 24-NOV-1997; 97US-0066770P.

PR 25-FEB-1998; 98US-0075945P.

PR 20-MAR-1998; 98US-0078910P.

PR 28-APR-1998; 98US-0083322P.

PR 07-MAY-1998; 98US-0084600P.

PR 28-MAY-1998; 98US-0087106P.

PR 02-JUN-1998; 98US-0087607P.

PR 02-JUN-1998; 98US-0087609P.

PR 02-JUN-1998; 98US-0087759P.

PR 03-JUN-1998; 98US-0088272P.

PR 04-JUN-1998; 98US-0088021P.

PR 04-JUN-1998; 98US-0088025P.

PR 04-JUN-1998; 98US-0088026P.

PR 04-JUN-1998; 98US-0088028P.

PR 04-JUN-1998; 98US-0088029P.

PR 04-JUN-1998; 98US-0088030P.

PR 04-JUN-1998; 98US-0088033P.

PR 04-JUN-1998; 98US-0088326P.

PR 05-JUN-1998; 98US-0088167P.

PR 05-JUN-1998; 98US-0088202P.

PR 05-JUN-1998; 98US-0088212P.

PR 05-JUN-1998; 98US-0088217P.

PR 09-JUN-1998; 98US-0088655P.

PR 10-JUN-1998; 98US-0088734P.

PR 10-JUN-1998; 98US-0088738P.

PR 10-JUN-1998; 98US-0088742P.

PR 10-JUN-1998; 98US-0088810P.

PR 10-JUN-1998; 98US-0088824P.

PR 10-JUN-1998; 98US-0088826P.

PR 11-JUN-1998; 98US-0088856P.

PR 11-JUN-1998; 98US-0088861P.

PR 12-JUN-1998; 98US-0088876P.

PR 12-JUN-1998; 98US-0089105P.

PR 16-JUN-1998; 98US-0089440P.

PR 16-JUN-1998; 98US-0089512P.

PR 16-JUN-1998; 98US-0089514P.

PR 17-JUN-1998; 98US-0089532P.

PR 17-JUN-1998; 98US-0089538P.

PR 17-JUN-1998; 98US-0089598P.

PR 17-JUN-1998; 98US-0089599P.

PR 17-JUN-1998; 98US-0089600P.

PR 17-JUN-1998; 98US-0089653P.

PR 18-JUN-1998; 98US-0089801P.

PR 18-JUN-1998; 98US-0089907P.

PR 18-JUN-1998; 98US-0089908P.

PR 19-JUN-1998; 98US-0089947P.

PR 19-JUN-1998; 98US-0089948P.

PR 19-JUN-1998; 98US-0089952P.

PR 22-JUN-1998; 98US-0090246P.

PR 22-JUN-1998; 98US-0090252P.

PR 22-JUN-1998; 98US-0090254P.

PR 23-JUN-1998; 98US-0090349P.

PR 23-JUN-1998; 98US-0090355P.

PR 24-JUN-1998; 98US-0090429P.

PR 24-JUN-1998; 98US-0090431P.

PR 24-JUN-1998; 98US-0090435P.

PR 24-JUN-1998; 98US-0090444P.

PR 24-JUN-1998; 98US-0090445P.

PR 24-JUN-1998; 98US-0090472P.

PR 24-JUN-1998; 98US-0090535P.

PR 24-JUN-1998; 98US-0090540P.

PR 24-JUN-1998; 98US-0090542P.

PR 24-JUN-1998; 98US-0090557P.

PR 25-JUN-1998; 98US-0090676P.

PR 25-JUN-1998; 98US-0090678P.

PR 25-JUN-1998; 98US-0090690P.

PR 25-JUN-1998; 98US-0090694P.

PR 25-JUN-1998; 98US-0090695P.

PR 25-JUN-1998; 98US-0090696P.

PR 26-JUN-1998; 98US-0090862P.

PR 26-JUN-1998; 98US-0090863P.

PR 01-JUL-1998; 98US-0091360P.

PR 01-JUL-1998; 98US-0091544P.

PR 02-JUL-1998; 98US-0091478P.

PR 02-JUL-1998; 98US-0091519P.

PR 02-JUL-1998; 98US-0091626P.

PR 02-JUL-1998; 98US-0091628P.

PR 02-JUL-1998; 98US-0091633P.

PR 02-JUL-1998; 98US-0091646P.

PR 02-JUL-1998; 98US-0091673P.

PR 07-JUL-1998; 98US-0091978P.

PR 07-JUL-1998; 98US-0091982P.

PR 09-JUL-1998; 98US-0092182P.

PR 10-JUL-1998; 98US-0092472P.

PR 20-JUL-1998; 98US-0093339P.

PR 30-JUL-1998; 98US-0094651P.

PR 04-AUG-1998; 98US-0095282P.



PR	14-SEP-1998;	98WO-US019094;
PR	14-SEP-1998;	98WO-US019177;
PR	16-SEP-1998;	98WO-US019330;
PR	16-SEP-1998;	98WO-US019437;
PR	07-OCT-1998;	98WO-US021141;
PR	29-OCT-1998;	98WO-US022991;
PR	29-OCT-1998;	98WO-US022992;
PR	20-NOV-1998;	98WO-US024855;
PR	01-DEC-1998;	98WO-US025108;
PR	05-JAN-1999;	98WO-US000106;
PR	08-MAR-1999;	98WO-US005028;
PR	10-MAR-1999;	98WO-US005190;
PR	20-APR-1999;	98WO-US008615;
PR	02-MAY-1999;	98WO-US010733;
PR	14-JUN-1999;	98WO-US012252;
PR	01-SEP-1999;	98WO-US020111;
PR	08-SEP-1999;	98WO-US020154;
PR	13-SEP-1999;	98WO-US020944;
PR	15-SEP-1999;	98WO-US021090;
PR	15-SEP-1999;	98WO-US021547;
PR	05-OCT-1999;	98WO-US023089;
PR	29-NOV-1999;	98WO-US028214;
PR	30-NOV-1999;	98WO-US028313;
PR	30-NOV-1999;	98WO-US028409;
PR	01-DEC-1999;	98WO-US028301;
PR	01-DEC-1999;	98WO-US028634;
PR	02-DEC-1999;	98WO-US028551;
PR	02-DEC-1999;	98WO-US028564;
PR	02-DEC-1999;	98WO-US028565;
PR	16-DEC-1999;	98WO-US030095;
PR	20-DEC-1999;	98WO-US030911;
PR	20-DEC-1999;	98WO-US030999;
PR	22-DEC-1999;	98WO-US030720;
PR	22-DEC-1999;	98WO-US031243;
PR	05-JAN-2000;	98WO-US031274;
PR	05-JAN-2000;	2000WO-US000219;
PR	06-JAN-2000;	2000WO-US000277;
PR	11-FEB-2000;	2000WO-US000376;
PR	16-FEB-2000;	2000WO-US003565;
PR	18-FEB-2000;	2000WO-US004341;
PR	18-FEB-2000;	2000WO-US004342;
PR	22-FEB-2000;	2000WO-US004414;
PR	24-FEB-2000;	2000WO-US004914;
PR	24-FEB-2000;	2000WO-US005004;
PR	01-MAR-2000;	2000WO-US005601;
PR	02-MAR-2000;	2000WO-US005746;
PR	02-MAR-2000;	2000WO-US005841;
PR	10-MAR-2000;	2000WO-US006319;
PR	15-MAR-2000;	2000WO-US006894;
PR	21-MAR-2000;	2000WO-US007377;
PR	21-MAR-2000;	2000WO-US007532;
PR	30-MAR-2000;	2000WO-US008439;
PR	17-MAY-2000;	2000WO-US013705;
PR	22-MAY-2000;	2000WO-US014042;
PR	02-JUN-2000;	2000WO-US014941;
PR	02-JUN-2000;	2000WO-US015264;
PR	28-JUL-2000;	2000WO-US020710;
PR	11-AUG-2000;	2000WO-US020231;
PR	23-AUG-2000;	2000WO-US023522;
PR	24-AUG-2000;	2000WO-US023328;
PR	08-NOV-2000;	2000WO-US030952;
PR	10-NOV-2000;	2000WO-US030873;
PR	10-DEC-2000;	2000WO-US032678;
PR	20-DEC-2000;	2000US-00747259;
PR	20-DEC-2000;	2000WO-US034956;
PR	28-DEC-2001;	2001US-00736494;
PR	28-FEB-2001;	2001WO-US006520;
PR	01-MAR-2001;	2001WO-US006666;
PR	09-MAR-2001;	2001US-00802706;
PR	14-MAR-2001;	2001US-00806899;
PR	22-MAR-2001;	2001US-00816744;
PR	05-APR-2001;	2001US-00828366;
PR	10-MAY-2001;	2001US-00854408;

PR	10-MAY-2001,	2001US-00854280.
PR	15-MAY-2001,	2001US-00860216.
PR	28-MAY-2001,	2001US-00860628.
PR	25-MAY-2001,	2001US-00866034.
PR	25-MAY-2001,	2001US-00866039.
PR	01-JUN-2001,	2001US-00872035.
PR	01-JUN-2001,	2001US-00872035.
PR	05-JUN-2001,	2001US-00874503.
PR	14-JUN-2001,	2001US-00882636.
PR	19-JUN-2001,	2001US-00886342.
PR	20-JUN-2001,	2001US-00891962.
PR	22-JUN-2001,	2001US-00887879.
PR	22-JUN-2001,	2001US-00920116.
PR	29-JUN-2001,	2001US-00921066.
PR	09-JUL-2001,	2001US-00921735.
PR	18-JUL-2001,	2001US-00908827.
PR	08-AUG-2001,	2001US-00924419.
PR	16-AUG-2001,	2001US-00927796.
PR	16-AUG-2001,	2001US-00931836.
PR	19-DEC-2001,	2001US-00928072.

(GETH ) GENENTECH INC.

AA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

WPI; 2003-341980/32.  
N-PSDB; ACD24080.

PT New secreted and transmembrane PRO nucleic acids, for treating  
 PT inflammation, organ failure, atherosclerosis, cardiac injury,  
 PT infertility, birth defects, premature aging, acquired immunodeficiency  
 PT syndrome (AIDS), or cancer.

Claim 12; Fig 456; 660pp; English.

The invention describes an isolated nucleic acid (I) comprising, or which has 80% sequence identity to, or the full-length coding sequence of, one or 75 nucleotide sequences, and which encodes a corresponding polypeptide selected from 275 amino acid sequences, where all sequences are given in the specification. The polypeptide encoded by (I) is used to detect PRO polypeptides, link a bioactive molecule to a cell expressing a PRO polypeptide, modulate a biological activity of a cell, stimulate a release of tumour necrosis factor (TNF)-alpha from human blood, modulate the uptake of glucose or free fatty acid by cells, stimulate or inhibit the proliferation or differentiation of cells or gene expression, stimulate the release of proteoglycans, stimulate the release of cytokine from peripheral blood mononuclear cells, inhibit the binding of A-peptide to factor VIIA, or detect the presence of tumour in a mammal. The nucleic acid and polypeptide encoded by it, are useful for treating inflammatory diseases, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, acquired immunodeficiency syndrome (AIDS), cancer, or diabetic complications. The nucleic acid is useful as hybridisation probes, in chromosome and gene mapping, and in generating antisense RNA or DNA. The polypeptides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. Both are useful in tissue typing. This is the amino acid sequence of a novel human secreted and transmembrane PRO polypeptide

```
Query Match      100.0%; Score 1608; DB 6; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.2e-120;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

1	Qy	1	MATARPMMWVLCALITALLGVTEHVLIANNVSCDHPSPNTVPSGSNQDILGAGAGEDARS	60
	D6	1	MATARPMMWVLCALITALLGVTEHVLIANNVSCDHPSPNTVPSGSNQDILGAGAGEDARS	60
	Qy	61	DDSSSRRIINGSDCDMHTQPQAAALLLRPNQLYCAGVLVHPQWLLTAAHCRKKVFRVRLGH	120
	D6	61	DDSSSRRIINGSDCDMHTQPQAAALLLRPNQLYCAGVLVHPQWLLTAAHCRKKVFRVRLGH	120



Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGSHNDLMLIKLNRIRPTKDVPRINVSCHPS 180  
|||||  
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGSHNDLMLIKLNRIRPTKDVPRINVSCHPS 180  
|||||  
Qy 181 AGTKCLVSGWTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDS 240  
|||||  
Db 181 AGTKCLVSGWTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDS 240  
|||||  
Qy 241 CQDGGGPPVVCNGSLQGLVSWGDYPCARPNGVYTNLCKFTKWIQTIQANS 293  
|||||  
Db 241 CQDGGGPPVVCNGSLQGLVSWGDYPCARPNGVYTNLCKFTKWIQTIQANS 293  
|||||  
RESULT 14  
ADA57425 standard; protein; 293 AA.  
ID ADA57425 standard; protein; 293 AA.  
XX AC ADA57425;  
XX AC  
DT 20-NOV-2003 (first entry)  
XX DE Human secreted protein #257.  
XX KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;  
XX KW cytosolic; cerebroprotective; neuroprotective; nootropic;  
XX KW cardiovascular; antiarteriosclerotic; gene therapy;  
XX KW human secreted protein; immune disorder; inflammation;  
XX KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
XX KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
XX KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
XX KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
XX KW triple helix formation; antisense gene therapy; forensic biology.  
XX OS Homo sapiens.  
XX PN WO2002102994-A2.  
XX PD 27-DEC-2002.  
XX PF 19-MAR-2002; 2002WO-US008278.  
XX PR 21-JUL-2001; 2001US-0277340P.  
XX PR 19-JUL-2001; 2001US-0306171P.  
XX PR 13-NOV-2001; 2001US-0331287P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Rosen CA, Ruben SM;  
XX DR WPI, 2003-167512/16.  
XX DR N-PSDB; ADA56532.  
XX PT New human secreted polypeptides and polynucleotides, useful for  
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory  
PT conditions, respiratory disorders, cancers, CNS disorders, or  
PT neurodegenerative disorders.  
XX FS Claim 13; SEQ ID NO 1618; 1754pp; English.  
XX CC The invention relates to 592 new human secreted polypeptides useful for  
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory  
CC conditions, respiratory disorders, cancers, CNS disorders, or  
CC neurodegenerative disorders, or polypeptides comprising an amino acid  
CC sequence at least 95% identical to the new sequences. The polypeptides,  
CC antibodies or antibody fragments that bind to the polypeptides, nucleic  
CC acids encoding the polypeptides, agonists or antagonists that binds to  
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical  
CC compositions for diagnosing, treating or preventing an e.g. immune  
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative

CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and  
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
CC polynucleotides are useful for chromosome identification, chromosome  
CC mapping, for controlling gene expression through triple helix formation  
CC or antisense DNA or RNA, in gene therapy, for identifying individuals  
CC from minute biological samples, in forensic biology, and as hybridization  
CC probes. The polypeptides are useful for as molecular weight markers on  
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)  
CC gels, to raise antibodies, for testing biological activities, and for  
CC treating or preventing neural disorders, immune system disorders,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal, proliferative and/or cancerous diseases. This sequence corresponds  
CC to one of the polypeptide of the invention. Note: The sequence data for  
CC this patent did form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 293 AA;

Query Match 100.0%; Score 1608; DB 6; Length 293;  
Best Local Similarity 100.0%; Pred. No. 4.2e-120;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
|||||  
Db 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
|||||  
Qy 61 DSSSRINSGDCDMHTQPHQALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120  
|||||  
Db 61 DSSSRINSGDCDMHTQPHQALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120  
|||||  
Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGSHNDLMLIKLNRIRPTKDVPRINVSCHPS 180  
|||||  
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGSHNDLMLIKLNRIRPTKDVPRINVSCHPS 180  
|||||  
Qy 181 AGTKCLVSGWTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDS 240  
|||||  
Db 181 AGTKCLVSGWTTKSPQVHPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAKGRDS 240  
|||||  
Qy 241 CQDGGGPPVVCNGSLQGLVSWGDYPCARPNGVYTNLCKFTKWIQTIQANS 293  
|||||  
Db 241 CQDGGGPPVVCNGSLQGLVSWGDYPCARPNGVYTNLCKFTKWIQTIQANS 293  
|||||

RESULT 15

ADA56974  
ID ADA56974 standard; protein; 293 AA.

XX AC ADA56974;

XX AC  
DT 20-NOV-2003 (first entry)

XX DE Human secreted protein #257.

XX KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;  
KW cytosolic; cerebroprotective; neuroprotective; nootropic;  
KW cardiovascular; antiarteriosclerotic; gene therapy;  
KW human secreted protein; immune disorder; inflammation;  
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;  
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;  
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;  
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;  
KW triple helix formation; antisense gene therapy; forensic biology.

XX OS Homo sapiens.

XX PN WO2002102994-A2.

XX PD 27-DEC-2002.

XX PF 19-MAR-2002; 2002WO-US008278.

XX PR 21-MAR-2001; 2001US-0277340P.

Search completed: July 4, 2005, 15:21:15  
Job time : 77 secs

PR 19-JUL-2001; 2001US-0306171P.  
PR 13-NOV-2001; 2001US-0331287P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX

PI Rosen CA, Ruben SM;  
XX  
XX WPI; 2003-167512/16.  
DR N-PSDB; ADA56078.  
XX

XX New human secreted polypeptides and polynucleotides, useful for  
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory  
PT conditions, respiratory disorders, cancers, CNS disorders, or  
PT neurodegenerative disorders.  
XX

PS Claim 13; SEQ ID NO 1164; 1754pp; English.  
XX

CC The invention relates to 592 new human secreted polypeptides useful for  
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory  
CC conditions, respiratory disorders, cancers, CNS disorders, or  
CC neurodegenerative disorders, or polypeptides comprising an amino acid  
CC sequence at least 93% identical to the new sequences. The polypeptides,  
CC antibodies or antibody fragments that bind to the polypeptides, nucleic  
CC acids encoding the polypeptides, agonists or antagonists that binds to  
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical  
CC compositions for diagnosing, treating or preventing an e.g. immune  
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,  
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and  
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders  
CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative  
CC disorders (e.g. Parkinson's disease or Alzheimer's disease) and  
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The  
CC polynucleotides are useful for chromosome identification, chromosome  
CC mapping, for controlling gene expression through triple helix formation  
CC or antisense DNA or RNA, in gene therapy, for identifying individuals  
CC from minute biological samples, in forensic biology, and as hybridization  
CC probes. The polypeptides are useful for as molecular weight markers on  
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)  
CC gels, to raise antibodies, for testing biological activities, and for  
CC treating or preventing neural disorders, immune system disorders,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal, proliferative and/or cancerous diseases. This sequence corresponds  
CC to one of the polypeptide of the invention. Note: The sequence data for  
CC this patent did form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 293 AA;

Query Match	100.0%	Score 1608;	DB 6;	Length 293;
Best Local Similarity	100.0%;	Pred. No. 4.2e-120;		
Matches 293;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MATARPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNODLGAGAGEDARS	60	
Db	1	MATARPMMWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSSNODLGAGAGEDARS	60	
Qy	61	DDSSRIINGSDCDMHTQFWQAALLRLPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH	120	
Db	61	DDSSRIINGSDCDMHTQFWQAALLRLPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH	120	
Qy	121	YLSLSPVYESGQMFQGVKSIHPHGYSHPGSHNDLMLIKLNRIRPTKDVRIINVSSHCPS	180	
Db	121	YLSLSPVYESGQMFQGVKSIHPHGYSHPGSHNDLMLIKLNRIRPTKDVRIINVSSHCPS	180	
Qy	181	AGTKCLVSGWGTTSQVHFPKVLQCLNIVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS	240	
Db	181	AGTKCLVSGWGTTSQVHFPKVLQCLNIVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS	240	
Qy	241	CQDSGGPVVNCNGSLQGLVSWGDFPCARPNGVYTNLCFTKWTQETIQANS	293	
Db	241	CQDSGGPVVNCNGSLQGLVSWGDFPCARPNGVYTNLCFTKWTQETIQANS	293	

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 01:15:51 ; Search time 24 Seconds  
(without alignments)  
911.341 Million cell updates/sec

Title: US-09-936-271C-14  
Perfect score: 1608  
Sequence: 1 MATARPPMMWVLCALITALL.....VYTNLCRFTKWIQETIQANS 293

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	293	4	US-09-509-908-2
2	1479	92.0	270	4	US-09-949-016-7712
3	1173	72.9	268	2	US-08-824-874-1
4	1173	72.9	268	3	US-09-210-084-1
5	1173	72.9	268	4	US-09-764-762-1
6	671	41.7	254	3	US-09-439-313-525
7	671	41.7	254	4	US-09-636-215-525
8	671	41.7	254	4	US-09-685-166A-525
9	671	41.7	254	4	US-09-679-426-525
10	671	41.7	254	4	US-09-759-143-525
11	671	41.7	254	4	US-09-651-236-525
12	666	41.4	254	3	US-09-439-313-523
13	666	41.4	254	4	US-09-636-215-523
14	666	41.4	254	4	US-09-685-166A-523
15	666	41.4	254	4	US-09-679-426-523
16	666	41.4	254	4	US-09-759-143-523
17	666	41.4	254	4	US-09-651-236-523
18	654.5	40.7	289	4	US-09-386-642-14
19	650.5	40.5	250	4	US-09-205-258-427
20	650.5	40.5	282	3	US-09-025-059-1
21	641.5	39.9	248	3	US-08-944-483-24
22	641.5	39.9	260	4	US-09-618-259-7
23	640	39.8	449	4	US-09-636-215-617
24	640	39.8	449	4	US-09-685-166A-617
25	640	39.8	449	4	US-09-679-426-617
26	640	39.8	449	4	US-09-759-143-617
27	640	39.8	449	4	US-09-651-236-617

28	639.5	39.8	260	3	US-09-008-271A-7	Sequence 7, Appli
29	638.5	39.7	260	3	US-09-070-526-2	Sequence 2, Appli
30	638.5	39.7	290	4	US-09-949-016-8166	Sequence 8166, Ap
31	638	39.7	248	1	US-08-744-026-1	Sequence 1, Appli
32	638	39.7	248	2	US-09-102-732-1	Sequence 1, Appli
33	638	39.7	248	3	US-09-261-767-1	Sequence 1, Appli
34	635	39.5	284	4	US-09-386-642-54	Sequence 54, Appli
35	632.5	39.3	288	4	US-09-386-642-13	Sequence 13, Appli
36	627	39.0	260	3	US-09-025-059-3	Sequence 3, Appli
37	627	39.0	260	4	US-09-618-259-8	Sequence 8, Appli
38	620	38.6	220	3	US-09-439-313-327	Sequence 327, App
39	620	38.6	220	3	US-09-352-616A-327	Sequence 327, App
40	620	38.6	220	4	US-09-232-149A-327	Sequence 327, App
41	620	38.6	220	4	US-09-636-215-327	Sequence 327, App
42	620	38.6	220	4	US-09-685-166A-327	Sequence 327, App
43	620	38.6	220	4	US-09-688-489-327	Sequence 327, App
44	620	38.6	220	4	US-09-679-426-327	Sequence 327, App
45	620	38.6	220	4	US-09-759-143-327	Sequence 327, App

ALIGNMENTS

RESULT 1  
US-09-509-908-2  
; Sequence 2, Application US/09509908  
; Patent No. 6589770  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company, N/A N/A  
; TITLE OF INVENTION: A Protease  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: T. David Reed  
; STREET: 5299 Spring Grove Avenue  
; CITY: Cincinnati  
; STATE: Ohio  
; COUNTRY: USA  
; ZIP: 45217-1087  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/509,908  
; FILING DATE: 28-Feb-2002  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Reed, T David  
; REGISTRATION NUMBER: 32,931  
; REFERENCE/DOCKET NUMBER: AA-264F  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 513-627-7025  
; TELEFAX: 513-627-6333  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 293 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-509-908-2

Query Match	100.0%;	Score 1608;	DB 4;	Length 293;
Best Local Similarity	100.0%;	Pred. No. 6e-163;		
Matches 293;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTVPSGNSODLGAGAGEDARS	60	
Db	1	MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHPSTVPSGNSODLGAGAGEDARS	60	
Qy	61	DDSSRLINGSDDCMHTQPWQAALLLPNOLYCCAVLVHPQWLLTAHCRKKVFRVLGH	120	

Db 61 DSSSRINGSDCDMHTQWQAALLRNQLYCGAVLVHPQWLLTAACRKKVFRVRLGH 120  
Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSCHPS 180  
Db 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSCHPS 180  
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKGRDS 240  
Db 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKGRDS 240  
Qy 241 CQSDSGPVVWVNGSLQGLVSGWDPVPCARPVRPGVYTNLCKFTKWIQETIQANS 293  
Db 241 CQSDSGPVVWVNGSLQGLVSGWDPVPCARPVRPGVYTNLCKFTKWIQETIQANS 293

RESULT 2  
US-09-949-016-7712  
; Sequence 7712, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7712  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-7712

Query Match 92.0%; Score 1479; DB 4; Length 270;  
Best Local Similarity 100.0%; Pred. No. 3e-149;  
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 EHVLANNDVSCDHPSTNTVPSSNODLGAGAGEDARSDDSSRIINGSDCDMHTQWQAAL 84  
Db 2 EHVLANNDVSCDHPSTNTVPSSNODLGAGAGEDARSDDSSRIINGSDCDMHTQWQAAL 61  
Qy 85 LLRNQLYCGAVLVHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPG 144  
Db 62 LLRNQLYCGAVLVHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPG 121  
Qy 145 YSHPGHNDMLIKLNRIRPTKDVRIINVSCHPSAGTKCLVSGWGTTKSPQVHPFKVL 204  
Db 122 YSHPGHNDMLIKLNRIRPTKDVRIINVSCHPSAGTKCLVSGWGTTKSPQVHPFKVL 181  
Qy 205 QCLNISVLSQKRCEDAYPRQIDDTMFCAGDKGRDS CQSDSGPVVWVNGSLQGLVSGWGDY 264  
Db 182 QCLNISVLSQKRCEDAYPRQIDDTMFCAGDKGRDS CQSDSGPVVWVNGSLQGLVSGWGDY 241  
Qy 265 PCARPVRPGVYTNLCKFTKWIQETIQANS 293  
Db 242 PCARPVRPGVYTNLCKFTKWIQETIQANS 270

RESULT 3  
US-08-824-874-1  
; Sequence 1, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN

; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 268 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: KERANOT02  
; CLONE: 820694  
US-08-824-874-1

Query Match 72.9%; Score 1173; DB 2; Length 268;  
Best Local Similarity 85.7%; Pred. No. 1.2e-116;  
Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;

Qy 38 PSNTVPSSGNSQ--DLGAGAGEDARSDDSSRIINGSDCDMHTQWQAALLRNQLYCGA 95  
Db 13 PLTECPGLGATRTWELGPKTPGRMTAAASMDPTAICT--PSRQAALLRNQLYCGA 70  
Qy 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPGYSHPGHNDLM 155  
Db 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQMFQGVKSIHPGYSHPGHNDLM 130  
Qy 156 LILNRRIRPTKDVRIINVSCHPSAGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQK 215  
Db 131 LILNRRIRPTKDVRIINVSCHPSAGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLSQK 190  
Qy 216 RCEDAYPRQIDDTMFCAGDKGRDS CQSDSGPVVWVNGSLQGLVSGWGDYPCARPVRPGVY 275  
Db 191 RCEDAYPRQIDDTMFCAGDKGRDS CQSDSGPVVWVNGSLQGLVSGWGDYPCARPVRPGVY 250  
Qy 276 TNLCKFTKWIQETIQANS 293  
Db 251 TNLCKFTKWIQETIQANS 268

RESULT 4  
US-09-210-084-1  
; Sequence 1, Application US/09210084  
; Patent No. 6197511  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/210,084

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/824,874

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0252 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: KERANOT02

CLONE: 820694

US-09-210-084-1

Query Match 72.9%; Score 1173; DB 3; Length 268;

Best Local Similarity 85.7%; Pred. No. 1.2e-116;

Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;

```
QY 38 PSNTVPSGSNQ--DLGAGAGEDARDSDSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGA 95
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 13 PLTPCPLGATRTWELGPKTGRMTAAASMDPTAICT--PSRQQAALLLRPNOLYCGA 70

QY 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQOMFQGVKSIPIHPGYSHPGHNSDLM 155
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQOMFQGVKSIPIHPGYSHPGHNSDLM 130

QY 156 LKLNRRIRPTKDVRRPINVSSHCPKSAGTKCLVSGWGTTKSPQVHPKVLQCLNISVLQK 215
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 131 LKLNRRIRPTKDVRRPINVSSHCPKSAGTKCLVSGWGTTKSPQVHPKVLQCLNISVLQK 190

QY 216 RCEDAYPRQIDDTMFCAGDKAGRSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 275
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 191 RCEDAYPRQIDDTMFCAGDKAGRSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 250

QY 276 TNLCKFTKWIQETIQANS 293
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 251 TNLCKFTKWIQETIQANS 268
```

#### RESULT 5

US-09-764-762-1

Sequence 1, Application US/09764762

Patent No. 6472195

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

INVENTOR: Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/764,762  
FILING DATE: 16-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/210,084  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0252 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KERANOT02  
CLONE: 820694  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-764-762-1

Query Match 72.9%; Score 1173; DB 4; Length 268;

Best Local Similarity 85.7%; Pred. No. 1.2e-116;

Matches 221; Conservative 8; Mismatches 25; Indels 4; Gaps 2;

```
QY 38 PSNTVPSGSNQ--DLGAGAGEDARDSDSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGA 95
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 13 PLTPCPLGATRTWELGPKTGRMTAAASMDPTAICT--PSRQQAALLLRPNOLYCGA 70

QY 96 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQOMFQGVKSIPIHPGYSHPGHNSDLM 155
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 71 VLHPQWLLTAACRKKVFRVRLGHYSLSPVYESGQOMFQGVKSIPIHPGYSHPGHNSDLM 130

QY 156 LKLNRRIRPTKDVRRPINVSSHCPKSAGTKCLVSGWGTTKSPQVHPKVLQCLNISVLQK 215
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 131 LKLNRRIRPTKDVRRPINVSSHCPKSAGTKCLVSGWGTTKSPQVHPKVLQCLNISVLQK 190

QY 216 RCEDAYPRQIDDTMFCAGDKAGRSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 275
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 191 RCEDAYPRQIDDTMFCAGDKAGRSCQDGGPVVCGSLQGLVSWGDPYPCARNRPGVY 250

QY 276 TNLCKFTKWIQETIQANS 293
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 251 TNLCKFTKWIQETIQANS 268
```

#### RESULT 6

US-09-439-313-525

Sequence 525, Application US/09439313

Patent No. 6329505

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Devin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yuqi

APPLICANT: Reed, Steven G.

APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary





Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGL 83  
Qy 121 YLSFVYESGQMFQGVKSIHPGSHGHNSDMLIKLNRRIRPTKDVPRINVSCHPCS 180  
Db 84 HSLEADQEPGSMQVEASLSVRHPEYNRPPLLNDMLIKLDESVSSEDTIRSIASQCPT 143  
Qy 181 AGTKLVSGWGTTSKSPQVHPFKVLOCLNISVLKRCEDAYPRQIDDTMFCAGDKAGRD 239  
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201  
Qy 240 SCQDGSQGVVVCNSLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 292  
Db 202 SCNGDSGGPLICNGYLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 254

## RESULT 9

US-09-679-426-525  
; Sequence 525, Application US/09679426  
; Patent No. 6759515

## GENERAL INFORMATION:

; APPLICANT: Xu Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C20  
; CURRENT APPLICATION NUMBER: US/09/679,426  
; CURRENT FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 895  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 525  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-679-426-525

Query Match 41.7%; Score 671; DB 4; Length 254;  
Best Local Similarity 45.7%; Pred. No. 3.4e-63;  
Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVDSCHPSNTVPSSGNDLQAGAGEDARS 60  
Db 1 MATAGNPMWFLGYLI-----AGSLVSG----- 26  
Qy 61 DDSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPQWVLSAAHCFQNSYTIIGL 120  
Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGL 83  
Qy 121 YLSFVYESGQMFQGVKSIHPGSHGHNSDMLIKLNRRIRPTKDVPRINVSCHPCS 180  
Db 84 HSLEADQEPGSMQVEASLSVRHPEYNRPPLLNDMLIKLDESVSSEDTIRSIASQCPT 143  
Qy 181 AGTKLVSGWGTTSKSPQVHPFKVLOCLNISVLKRCEDAYPRQIDDTMFCAGDKAGRD 239  
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201  
Qy 240 SCQDGSQGVVVCNSLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 292  
Db 202 SCNGDSGGPLICNGYLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 254

## RESULT 10

US-09-759-143-525  
; Sequence 525, Application US/09759143  
; Patent No. 680746

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C23  
; CURRENT APPLICATION NUMBER: US/09/759,143  
; CURRENT FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 934  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 525  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-759-143-525

Query Match 41.7%; Score 671; DB 4; Length 254;  
Best Local Similarity 45.7%; Pred. No. 3.4e-63;  
Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVDSCHPSNTVPSSGNDLQAGAGEDARS 60  
Db 1 MATAGNPMWFLGYLI-----AGSLVSG----- 26  
Qy 61 DDSSRIINGSDCDMHTOPWQAALLRPNOLYCGAVLVHPQWVLSAAHCFQNSYTIIGL 120  
Db 27 --SCSIIINGEDCSPHSQPWAALVME-NELFCSGVLVHPQWVLSAAHCFQNSYTIIGL 83  
Qy 121 YLSFVYESGQMFQGVKSIHPGSHGHNSDMLIKLNRRIRPTKDVPRINVSCHPCS 180  
Db 84 HSLEADQEPGSMQVEASLSVRHPEYNRPPLLNDMLIKLDESVSSEDTIRSIASQCPT 143  
Qy 181 AGTKLVSGWGTTSKSPQVHPFKVLOCLNISVLKRCEDAYPRQIDDTMFCAGDKAGRD 239  
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLDPLVHPFMSFCAAGGQDQKD 201  
Qy 240 SCQDGSQGVVVCNSLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 292  
Db 202 SCNGDSGGPLICNGYLQSLVSGWGDYPCARPNNRPGVYTNLCKFTKWIOTIOAN 254

## RESULT 11

US-09-651-236-525  
; Sequence 525, Application US/09651236  
; Patent No. 6818751

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.

```
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651,236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 525
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapien
;
US-09-651-236-525

Query Match 41.7%; Score 671; DB 4; Length 254;
Best Local Similarity 45.7%; Pred. No. 3.4e-63;
Matches 134; Conservative 47; Mismatches 72; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPGSNQDLGAGAGEDARS 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MATAGNPGWFLGYLI-----LGV-----AGSLVSG----- 26

Qy 61 DSSSRIINGSDCDMHTOPWQAALLRPOLYCGAVLVHPQWLTAACRKKVFRVRLGH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 --SCSQIINGEDCSPHSQWQAALVME--NELFCGVLVHPQWVLSATHCFQNSYTIIGLGL 83

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNNRIRPTKDVRRPINVSSHCP 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 HSLADQEPGQMVESASLSVRHPEYNNRPLLANDMLIKLDESVSSESDTIRISIASQCPT 143

Qy 181 AGTKCLVSGWGTTSKSPQVHPKVLQCLNISVLQKRCEDAYPROIDDTMFCA-GDKAGRD 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLYDPLYHPSMFCAGGGQDQKD 201

Qy 240 SCQDGGGPPVVCNGSLQGLVSGWGDYPCARPNNRPGVYTNLCKFTWKIQTIOAN 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 SCNGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254

RESULT 12
US-09-439-313-523
; Sequence 523, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 523
```

```
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = Any amino acid
;
US-09-439-313-523

Query Match 41.4%; Score 666; DB 3; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.2e-62;
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPGSNQDLGAGAGEDARS 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MATAGNPGWFLGYLI-----LGV-----AGSLVSG----- 26

Qy 61 DSSSRIINGSDCDMHTOPWQAALLRPOLYCGAVLVHPQWLTAACRKKVFRVRLGH 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 --SCSQIINGEDCSPHSQWQAALVME--NELFCGVLVHPQWVLSATHCFQNSYTIIGLGL 83

Qy 121 YSLSPVYESGQMFQGVKSIHPGYSHPGHNSDMLIKLNNRIRPTKDVRRPINVSSHCP 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 HSLADQEPGQMVESASLSVRHPEYNNRPLLANDMLIKLDESVSSESDTIRISIASQCPT 143

Qy 181 AGTKCLVSGWGTTSKSPQVHPKVLQCLNISVLQKRCEDAYPROIDDTMFCA-GDKAGRD 239
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLYDPLYHPSMFCAGGGQDQXD 201

Qy 240 SCQDGGGPPVVCNGSLQGLVSGWGDYPCARPNNRPGVYTNLCKFTWKIQTIOAN 292
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 202 SCNGDSGGPLICNGYLQGLVSGFKAPCGQGVPGVYTNLCKFTWIEKTVQAS 254

RESULT 13
US-09-636-215-523
; Sequence 523, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 523
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; NAME/KEY: VARIANT
; LOCATION: (1)...(254)
; OTHER INFORMATION: Xaa = any amino acid
```

US-09-636-215-523

Query Match 41.4%; Score 666; DB 4; Length 254;  
Best Local Similarity 45.4%; Pred. No. 1.2e-62;  
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPSVYESGQMGVVKSIHPHGYSHPGHNDLMLIKLNRRIPTKDVRRPINVSSHCP 180  
DB 84 HSLADQEPGQMVESLVSVRHPEYRNPRLANDLMLIKLDSVESDSTIRISIASQCPT 143

QY 181 AGTKLVSGMTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCA-GDKAGRD 239  
DB 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVVSEVCSKLYDPLVHPSMFCAGGQXQXD 201

QY 240 SCQDGSQPVVNCVSLQSLGWSGDYPCARPNRPGVYTNLCKFTKWIQETIOAN 292  
DB 202 SCNGDSGGPLICNGYLOGLVSGFKAPCGQGVGVYTNLCKFTWIEKTVQAS 254

## RESULT 14

US-09-685-166A-523  
; Sequence 523, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C21  
; CURRENT APPLICATION NUMBER: US/09/685.166A  
; CURRENT FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 898  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 523  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Made in a lab  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(254)  
; OTHER INFORMATION: Xaa = 'any amino acid'  
US-09-685-166A-523

Query Match 41.4%; Score 666; DB 4; Length 254;  
Best Local Similarity 45.4%; Pred. No. 1.2e-62;  
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPSVYESGQMGVVKSIHPHGYSHPGHNDLMLIKLNRRIPTKDVRRPINVSSHCP 180  
DB 84 HSLADQEPGQMVESLVSVRHPEYRNPRLANDLMLIKLDSVESDSTIRISIASQCPT 143

QY 181 AGTKLVSGMTTKSPQVHFVKVQLCLNISVLQKRCEDAYPRQIDDTMFCA-GDKAGRD 239  
DB 144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVVSEVCSKLYDPLVHPSMFCAGGQXQXD 201

QY 240 SCQDGSQPVVNCVSLQSLGWSGDYPCARPNRPGVYTNLCKFTKWIQETIOAN 292  
DB 202 SCNGDSGGPLICNGYLOGLVSGFKAPCGQGVGVYTNLCKFTWIEKTVQAS 254

## RESULT 15

US-09-679-426-523  
; Sequence 523, Application US/09679426  
; Patent No. 6759515  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqui  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Li, Samuel  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Hepler, William  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.427C20  
; CURRENT APPLICATION NUMBER: US/09/679.426  
; CURRENT FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 895  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 523  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Made in a lab  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(254)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-679-426-523

Query Match 41.4%; Score 666; DB 4; Length 254;  
Best Local Similarity 45.4%; Pred. No. 1.2e-62;  
Matches 133; Conservative 46; Mismatches 74; Indels 40; Gaps 6;

QY 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTVPSSNQDLGAGAGEDARS 60  
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

QY 61 DDSSRIINGSDCDMHTOPWQAALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
DB 27 --SCSQIINGDCSPHSQPWQAALVME-NELFCGVLVHPQWLVLSATHCFQNSYITIGLGL 83

QY 121 YLSLPSVYESGQMGVVKSIHPHGYSHPGHNDLMLIKLNRRIPTKDVRRPINVSSHCP 180  
DB 1 MATAGNPMGWFLGYLI-----LGV-----AGSLVSG----- 26

```
Db      84 HSLEADQEPGSQMVESLSVRHPEYNRPLLANDLMLIKLDESVSSESDTIRSIASQCPT 143
Qy      181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVL SOKRCEDAYPRQ1DDTMFCA-GDKAGRD 239
Db      144 AGNSCLVSGWGLLANGRM--PTVLQCVNVSVSEVCSKLYDPLYHPSMFCAGGGQXQXD 201
Qy      240 SCQGDGSGPVVNCNGSLOGLVSMGDYPFCARNRPGVYTNLCKFTKWIQETIQAN 292
Db      202 SCNGDSGGPLICNGYLQGLVSFGKAPCGQGVGPGVYTNLCKFTWIEKTVQAS 254
```

Search completed: July 4, 2005, 15:07:26  
Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 09:55:49 ; Search time 400 Seconds  
(without alignments)  
282.595 Million cell updates/sec

Title: US-09-936-271c-14  
Perfect score: 1608  
Sequence: 1 MATARPPMMVLCALITALL.....VYTNLCRFTKWIQTIQANS 293

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1722976 seqs, 385795295 residues

Total number of hits satisfying chosen parameters: 1722976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	293	9	US-09-739-907-82
2	1608	100.0	293	9	US-09-739-907-97
3	1608	100.0	293	9	US-09-989-722-309
4	1608	100.0	293	9	US-09-989-723-309
5	1608	100.0	293	9	US-09-989-729-309
6	1608	100.0	293	9	US-09-989-727-309
7	1608	100.0	293	9	US-09-989-731-309
8	1608	100.0	293	9	US-09-989-732-309
9	1608	100.0	293	9	US-09-991-073-309
10	1608	100.0	293	9	US-09-990-442-309
11	1608	100.0	293	9	US-09-991-163-309

12	1608	100.0	293	9	US-09-993-604-309
13	1608	100.0	293	9	US-09-990-456-309
14	1608	100.0	293	9	US-09-989-721-309
15	1608	100.0	293	9	US-09-992-598-309
16	1608	100.0	293	9	US-09-989-293A-309
17	1608	100.0	293	9	US-09-989-735-309
18	1608	100.0	293	9	US-09-990-444-309
19	1608	100.0	293	9	US-09-991-181-309
20	1608	100.0	293	9	US-09-989-730-309
21	1608	100.0	293	9	US-09-990-436-309
22	1608	100.0	293	9	US-09-993-687-309
23	1608	100.0	293	10	US-09-989-734-309
24	1608	100.0	293	10	US-09-997-653-309
25	1608	100.0	293	10	US-09-989-724-309
26	1608	100.0	293	10	US-09-989-728-309
27	1608	100.0	293	10	US-09-990-441-309
28	1608	100.0	293	10	US-09-993-667-309
29	1608	100.0	293	10	US-09-997-428-309
30	1608	100.0	293	10	US-09-997-666-309
31	1608	100.0	293	10	US-09-990-438-309
32	1608	100.0	293	10	US-09-990-562-309
33	1608	100.0	293	10	US-09-990-711-309
34	1608	100.0	293	10	US-09-989-726-309
35	1608	100.0	293	10	US-09-989-156-309
36	1608	100.0	293	10	US-09-990-437-309
37	1608	100.0	293	10	US-09-991-157-309
38	1608	100.0	293	10	US-09-997-514-309
39	1608	100.0	293	10	US-09-997-573-309
40	1608	100.0	293	10	US-09-991-172-309
41	1608	100.0	293	10	US-09-990-726-309
42	1608	100.0	293	10	US-09-997-559-309
43	1608	100.0	293	10	US-09-997-601-309
44	1608	100.0	293	10	US-09-990-443-309
45	1608	100.0	293	10	US-09-991-854-309

#### ALIGNMENTS

RESULT 1  
US-03-739-907-82  
; Sequence 82, Application US/09739907  
; Patent No. US20010012889A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 36 Human Secreted Proteins  
; FILE REFERENCE: PZ022P1  
; CURRENT APPLICATION NUMBER: US/09/739,907  
; CURRENT FILING DATE: 2000-12-20  
; PRIOR APPLICATION NUMBER: 09/348,457  
; PRIOR FILING DATE: 1999-07-07  
; PRIOR APPLICATION NUMBER: 60/070,567  
; PRIOR FILING DATE: 1998-01-07  
; PRIOR APPLICATION NUMBER: 60/070,692  
; PRIOR FILING DATE: 1998-01-07  
; PRIOR APPLICATION NUMBER: 60/070,704  
; PRIOR FILING DATE: 1998-01-07  
; PRIOR APPLICATION NUMBER: 60/070,658  
; PRIOR FILING DATE: 1998-01-07  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 82  
; LENGTH: 293  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-739-907-82

Query Match 100.0%; Score 1608; DB 9; Length 293;  
Best Local Similarity 100.0%; Pred. No. 9.1e-150;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTNTPVPGSNODLCAGAGEDARS 60  
|||||

Db 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTNTVPSGSNQDLGAGGEDARS 60

Qy 61 DSSSRRIINGSDCMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120

Db 61 DSSSRRIINGSDCMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMGQGVKSIHPHGYSHPGHNSNDMLIKLNRRIRPTKDVPRPINVSSHCP 180

Db 121 YSLSPVYESGQMGQGVKSIHPHGYSHPGHNSNDMLIKLNRRIRPTKDVPRPINVSSHCP 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVLCQLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240

Db 181 AGTKCLVSGWGTTKSPQVHFVKVLCQLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240

Qy 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

Db 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

RESULT 2

US-09-739-907-97

; Sequence 97, Application US/09733907

; Patent No. US20010012889A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 36 Human Secreted Proteins

; FILE REFERENCE: P2022P1

; CURRENT APPLICATION NUMBER: US/09/739,907

; CURRENT FILING DATE: 2000-12-20

; PRIOR APPLICATION NUMBER: 09/348,457

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: 60/070,567

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,692

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,704

; PRIOR FILING DATE: 1998-01-07

; PRIOR APPLICATION NUMBER: 60/070,658

; PRIOR FILING DATE: 1998-01-07

; NUMBER OF SEQ ID NOS: 196

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 97

; LENGTH: 293

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-739-907-97

Query Match 100.0%; Score 1608; DB 9; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.1e-150;

Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTNTVPSGSNQDLGAGGEDARS 60

Db 1 MATARPPMMVLCALITALLIGVTEHVLANNVSCDHPSTNTVPSGSNQDLGAGGEDARS 60

Qy 61 DSSSRRIINGSDCMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120

Db 61 DSSSRRIINGSDCMHTQPMQAAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMGQGVKSIHPHGYSHPGHNSNDMLIKLNRRIRPTKDVPRPINVSSHCP 180

Db 121 YSLSPVYESGQMGQGVKSIHPHGYSHPGHNSNDMLIKLNRRIRPTKDVPRPINVSSHCP 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVLCQLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240

Db 181 AGTKCLVSGWGTTKSPQVHFVKVLCQLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240

Qy 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

Db 241 CQDSGGPVVNCVCSLQGLVSGWGDYPCARPNGPVYTNLCCKTKWIOETIQANS 293

RESULT 3

US-09-989-722-309

; Sequence 309, Application US/09989722

; Patent No. US20020072067A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Baker, Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Fong, Sherman

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Kijavini, Ivar J.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas P.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2730P1C63

; CURRENT APPLICATION NUMBER: US/09/989,722

; CURRENT FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/049787

; PRIOR FILING DATE: 1997-06-16

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/065186

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066770

; PRIOR FILING DATE: 1997-11-24

; PRIOR APPLICATION NUMBER: 60/075945

; PRIOR FILING DATE: 1998-02-25

; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/083322

; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/087106

; PRIOR FILING DATE: 1998-05-28

; PRIOR APPLICATION NUMBER: 60/087607

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087759

; PRIOR FILING DATE: 1998-06-02

; PRIOR APPLICATION NUMBER: 60/087827

; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088025

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088026

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088028

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088029

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088030

; PRIOR FILING DATE: 1998-06-04

; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355

; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;  
Best Local Similarity 100.0%; Pred. No. 9.1e-150;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSNTVPSCNQDLGAGAGEDARS 60  
Db 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSNTVPSCNQDLGAGAGEDARS 60  
Qy 61 DSSSRIINGSDCDMHTQPMQALLLRPNOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120  
Db 61 DSSSRIINGSDCDMHTQPMQALLLRPNOLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120  
Qy 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRIPTKDVRIPIVSSHCP 180  
Db 121 YLSLSPVYESGQMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRIPTKDVRIPIVSSHCP 180  
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSQKRCEDAYPRQIDDTMFCAGDKAGRDS 240



```
Db 181 AGTKCLVSGWGTTSKQVHPFKVLQCLNISVLQKRCEDAYPRQIDDTMFCAQDKAGRDS 240
Qy 241 CQGDGGPVVVCNGLQGLVSGMDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
Db 241 CQGDGGPVVVCNGLQGLVSGMDYPCARPNRPGVYTNLCKFTKWIQETIQANS 293

RESULT 4
US-09-989-723-309
; Sequence 309, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC82
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
```

```

; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHSNTVPSGNODLGGAGEDARS 60
DB 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHSNTVPSGNODLGGAGEDARS 60

```

```

QY 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCKKVRVLGH 120
DB 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCKKVRVLGH 120
QY 121 YSLSPVYESGQMFQGVKSIPIHPGYSHPGHSNDLMLIKLNRRIPTKDVREINVSSHCP 180
DB 121 YSLSPVYESGQMFQGVKSIPIHPGYSHPGHSNDLMLIKLNRRIPTKDVREINVSSHCP 180
QY 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSOKRCEDAYPRQIDDTMFCAGDKAGRDS 240
DB 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSOKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQGDGGPVVVCNGSLQGLVSGMDYPCARNRPGVYVTLCKFTKWIQETIOANS 293
DB 241 CQGDGGPVVVCNGSLQGLVSGMDYPCARNRPGVYVTLCKFTKWIQETIOANS 293

RESULT 5
US-09-989-279-309
; Sequence 309, Application US/09989279
; Patent No. US20020072496A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C56
; CURRENT APPLICATION NUMBER: US/09/989,279
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02

```

PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18	PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19	PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22	PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23	PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24	PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25	PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26	PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26	PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01	PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02	PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01	PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02	PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02	PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02	PRIOR APPLICATION NUMBER: 60/091798
PRIOR FILING DATE: 1998-07-07	PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07	PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-07	PRIOR APPLICATION NUMBER: 60/092182

[illegible]

## RESULT 6

```

US-09-989-727-309
; Sequence 309, Application US/09989727
; Patent No. US20020072497A1
GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC65
; CURRENT APPLICATION NUMBER: US/09/989,727
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910

```

; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626

; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;  
Best Local Similarity 100.0%; Pred. No. 9.1e-150; Indels 0; Gaps 0;  
Matches 293; Conservative 0; Mismatches 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLIANNVSCDHPSTNTPSGSNQDLGAGAGEDARS 60  
Db 1 MATARPPMMVLCALITALLGVTEHVLIANNVSCDHPSTNTPSGSNQDLGAGAGEDARS 60  
Qy 61 DSSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
Db 61 DSSSRIINGSDCDMHTQPMQAAALLLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
Qy 121 YSLSPVYESGOQMPQGVKSTPHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSCHPS 180  
Db 121 YSLSPVYESGOQMPQGVKSTPHPGYSHPGHNDMLIKLNRIRPTKDVRIINVSCHPS 180  
Qy 181 AGTKCLVSGWGTTKSPQVHFPPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
Db 181 AGTKCLVSGWGTTKSPQVHFPPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
Qy 241 CQDSDGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS 293  
Db 241 CQDSDGPPVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCCKFTKWIQETIQANS 293

## RESULT 7

US-09-989-731-309  
; Sequence 309, Application US/09989731  
; Patent No. US20020103125A1

## GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: KJavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730P1C70

; CURRENT APPLICATION NUMBER: US/09/989,731  
; CURRENT FILING DATE: 2001-11-20  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17

[illegible]

```
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match      100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRINSGDCDMHTQWQAALLLRPNQYLCGAVLVHPQWLLTAHCRKKVFRVLGH 120
Db 61 DSSSRINSGDCDMHTQWQAALLLRPNQYLCGAVLVHPQWLLTAHCRKKVFRVLGH 120

Qy 121 YLSPPVYESGQMGQGVKSIHPHGVSHPGHNDMLKLNRRIRPTKDVRPINVSSHCP 180
Db 121 YLSPPVYESGQMGQGVKSIHPHGVSHPGHNDMLKLNRRIRPTKDVRPINVSSHCP 180

Qy 181 AGTKCLVSGWMTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRDIDTDFMFCAGDKGRDS 240
Db 181 AGTKCLVSGWMTTKSPQVHFVKVLQCLNISVLQKRCEDAYPRDIDTDFMFCAGDKGRDS 240

Qy 241 CQDGGGPPVVCNGLQGLVSGDYPFCARPNRPGVYTNLCCKTKWIOETIQANS 293
Db 241 CQDGGGPPVVCNGLQGLVSGDYPFCARPNRPGVYTNLCCKTKWIOETIQANS 293

RESULT 8
US-09-989-732-309
; Sequence 309, Application US/09989732
; Patent No. US20020123463A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
```

```
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC57
; CURRENT APPLICATION NUMBER: US/09/989, 732
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
```



/	PRIOR FILING DATE: 1998-06-25
/	PRIOR APPLICATION NUMBER: 60/090695
/	PRIOR FILING DATE: 1998-06-25
/	PRIOR APPLICATION NUMBER: 60/090696
/	PRIOR FILING DATE: 1998-06-25
/	PRIOR APPLICATION NUMBER: 60/090862
/	PRIOR FILING DATE: 1998-06-26
/	PRIOR APPLICATION NUMBER: 60/090863
/	PRIOR FILING DATE: 1998-06-26
/	PRIOR APPLICATION NUMBER: 60/091360
/	PRIOR FILING DATE: 1998-07-01
/	PRIOR APPLICATION NUMBER: 60/091478
/	PRIOR FILING DATE: 1998-07-02
/	PRIOR APPLICATION NUMBER: 60/091544
/	PRIOR FILING DATE: 1998-07-01
/	PRIOR APPLICATION NUMBER: 60/091519
/	PRIOR FILING DATE: 1998-07-02
/	PRIOR APPLICATION NUMBER: 60/091626
/	PRIOR FILING DATE: 1998-07-02
/	PRIOR APPLICATION NUMBER: 60/091633
/	PRIOR FILING DATE: 1998-07-02
/	PRIOR APPLICATION NUMBER: 60/091978
/	PRIOR FILING DATE: 1998-07-07
/	PRIOR APPLICATION NUMBER: 60/091982
/	PRIOR FILING DATE: 1998-07-07
/	PRIOR APPLICATION NUMBER: 60/092182
/	PRIOR FILING DATE: 1998-07-09
Query Match 100.0%; Score 1608; DB 9; Length 293;	
Best Local Similarity 100.0%; Pred. No. 9.le=150;	
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTPSGSNQDLGAGAGEDARS 60
Db	1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTPSGSNQDLGAGAGEDARS 60
QY	61 DSSSRIINGSDCDMHTQPWOAALLLRPNOLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120
Db	61 DSSSRIINGSDCDMHTQPWOAALLLRPNOLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120
QY	121 YSLSPVYESQQOMFGVKSTPHPGYSHPGHSNDMLIKLNRRIRPTKDVRPINVSCHPS 180
Db	121 YSLSPVYESQQOMFGVKSTPHPGYSHPGHSNDMLIKLNRRIRPTKDVRPINVSCHPS 180
QY	181 AGTKCLVSGWGTTKSPOVHPEPKVLQCLNISVLQKRCEADAYPROIDDTMFCAGDKAGRDS 240
Db	181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEADAYPROIDDTMFCAGDKAGRDS 240
QY	241 CQDGSGBPVCNGSLQGLSVGWGDYPCCARPNRPGVYTNLCKFTKWIQETIOANS 293
Db	241 CQDGSGBPVCNGSLQGLSVGWGDYPCCARPNRPGVYTNLCKFTKWIQETIOANS 293
RESULT 9	
US-09-991-073-309	
Sequence 309, Application US/09991073	
Patent No. US20020127576A1	
GENERAL INFORMATION:	
APPLICANT: Ashkenazi, Avi J.	
APPLICANT: Baker, Kevin P.	
APPLICANT: Botstein, David	
APPLICANT: Desnoyers, Luc	
APPLICANT: Eaton, Dan L.	
APPLICANT: Ferrara, Napoleone	
APPLICANT: Fong, Sherman	
APPLICANT: Gerber, Hanspeter	
APPLICANT: Gerritsen, Mary E.	
APPLICANT: Goddard, Audrey	
APPLICANT: Godowski, Paul J.	
APPLICANT: Grimaldi, J. Christopher	
APPLICANT: Gurney, Austin L.	
APPLICANT: Kijavlin, Ivar J.	
APPLICANT: Naquier, Mary A.	

APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Scrated and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC15  
CURRENT APPLICATION NUMBER: US/09/991.073  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/045787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/065770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24

;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;  
Best Local Similarity 100.0%; Pred. No. 9.1e-150;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60  
Db 1 MATARPPWVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRINSGDCMHTQWQAALLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120  
Db 61 DSSSRINSGDCMHTQWQAALLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGOOMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRIPIVNSSHCPS 180  
Db 121 YSLSPVYESGOOMFQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVRIPIVNSSHCPS 180

Qy 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPROIDDTMFCAKGRDS 240  
Db 181 AGTKCLVSGWGTTKSPQVHFVKVQLCLNISVLQKRCEDAYPROIDDTMFCAKGRDS 240

Qy 241 CQDGGGPPVWCVNGSLQGLVSGDYPFCARPNRPVYTNLCCKTKVIQETIQANS 293  
Db 241 CQDGGGPPVWCVNGSLQGLVSGDYPFCARPNRPVYTNLCCKTKVIQETIQANS 293

## RESULT 10

US-09-930-442-309  
Sequence 309, Application US/09990442  
Patent No. US20020132252A1

## GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi J.  
;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnovers, Luc  
;; APPLICANT: Eaton, Dan L.  
;; APPLICANT: Ferrara, Napoleone

;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730P1C8  
CURRENT APPLICATION NUMBER: US/09/990,442  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05

; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445

; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.1e-150; Mismatches 0; Indels 0; Gaps 0;

Matches 293; Conservative 0;

Qy	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNODLGAGEDARS	60
Db	1	MATARPMMVLCALITALLGVTEHVLANNVSCDHPSTNTPVSGSNODLGAGEDARS	60
Qy	61	DDSSRIINGSDCDMHTQPMQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH	120
Db	61	DDSSRIINGSDCDMHTQPMQAALLLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVLGH	120
Qy	121	YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVPRINVSCHPS	180
Db	121	YSLSPVYESGQMPQGVKSIHPGYSHPGHSNDLMLIKLNRIRPTKDVPRINVSCHPS	180
Qy	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRQDDTMFCAGDKAGRDS	240
Db	181	AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQKRCEDAYPRQDDTMFCAGDKAGRDS	240
Qy	241	CQGDGGPVVVCNGLQGLVSWGDYPCARPNNRPGVYTNLCKFTKWIQETIQANS	293
Db	241	CQGDGGPVVVCNGLQGLVSWGDYPCARPNNRPGVYTNLCKFTKWIQETIQANS	293

Sequence 309, Application US/09991163  
Patent No. US20020132253A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC17  
CURRENT APPLICATION NUMBER: US/09/991,163  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23

```

; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLGVTEHVLNNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60
Db 1 MATARPPMMVLCALITALLGVTEHVLNNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DSSSRIINGSDCMHTQWQAALLRRNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120
Db 61 DSSSRIINGSDCMHTQWQAALLRRNQLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMFQGVKSIHPHGYSHPHGNSDMLIKLNRIIRPTKDVRPINVSSHCP 180
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPHGNSDMLIKLNRIIRPTKDVRPINVSSHCP 180

Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQSKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNISVLQSKRCEDAYPRQIDDTMFCAGDKAGRDS 240

; Sequence 309, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavins, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04

RESULT 12
US-09-993-604-309
; Sequence 309, Application US/09993604
; Patent No. US20020137075A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavins, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C25
; CURRENT APPLICATION NUMBER: US/09/993,604
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
```

; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089948  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/089952  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090246

; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090252  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090254  
; PRIOR FILING DATE: 1998-06-22  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090355  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;

Best Local Similarity 100.0%; Pred. No. 9.1e-150;

Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Db 1 MATARPPMMVLCALITALLIGVTEHVLANNDVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60

Qy 61 DDSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAACHRKKVFRVLGH 120

Db 61 DSSSRINGSDCMHTOPWQAALLRPNQLYCCAVLPHQWLLTAHCRKKVFRVLGH 120  
Qy 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLKLNRRIRPTKDVREINVSCHPS 180  
Db 121 YSLSPVYESGQMFQGVKSIHPHGYSHPGHNDLMLKLNRRIRPTKDVREINVSCHPS 180  
Qy 181 AGTKLVSGWGTTSKSPQVHFVKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240  
Db 181 AGTKLVSGWGTTSKSPQVHFVKVLQCLNISVLQKRCEDAYPRQIDDTMFCAKGRDS 240  
Qy 241 CQDSGGPWCNGLSGVLVSGDYPFCARPNRPVYTNLCKTKWIOETIQANS 293  
Db 241 CQDSGGPWCNGLSGVLVSGDYPFCARPNRPVYTNLCKTKWIOETIQANS 293

RESULT 13

US-09-930-456-309  
; Sequence 309, Application US/09990456  
; Patent No. US20020137890A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnovers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2730FIC22  
; CURRENT APPLICATION NUMBER: US/09/990,456  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609

; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-03  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18



```

Query Match          100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATARPPWMMVLCALITALLGVTEHVLANNVSCDHPSNTVPSGNSQDLGAGAGEDARS 60
Db 1 MATARPPWMMVLCALITALLGVTEHVLANNVSCDHPSNTVPSGNSQDLGAGAGEDARS 60
Qy 61 DSSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Db 61 DSSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGH 120
Qy 121 YSLSPVYESGQMFQGVKSIPIHPCGYSHPGHSNDLMLIKLNRRIPTKDVPRPINVSSHCPS 180
Db 121 YSLSPVYESGQMFQGVKSIPIHPCGYSHPGHSNDLMLIKLNRRIPTKDVPRPINVSSHCPS 180
Qy 181 AGTKCLVSGWGTTKSPVHPFKVLQCLNLSVLQKRCEDAYPROIDDTMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTKSPVHPFKVLQCLNLSVLQKRCEDAYPROIDDTMFCAGDKAGRDS 240
Qy 241 CQDGSQGPVVVCGNSLQGLVSGWGYPCARPNRPGVYTNLCKFTKWIQETIQANS 293
Db 241 CQDGSQGPVVVCGNSLQGLVSGWGYPCARPNRPGVYTNLCKFTKWIQETIQANS 293

RESULT 14
US-09-989-721-309
; Sequence 309, Application US/09989721
; Patent No. US20020142961A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C55
; CURRENT APPLICATION NUMBER: US/09/989,721
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

```

1 PRIOR APPLICATION NUMBER: 60/083322  
2 PRIOR FILING DATE: 1998-04-28  
3 PRIOR APPLICATION NUMBER: 60/084600  
4 PRIOR FILING DATE: 1998-05-07  
5 PRIOR APPLICATION NUMBER: 60/087106  
6 PRIOR FILING DATE: 1998-05-28  
7 PRIOR APPLICATION NUMBER: 60/087607  
8 PRIOR FILING DATE: 1998-06-02  
9 PRIOR APPLICATION NUMBER: 60/087609  
10 PRIOR FILING DATE: 1998-06-02  
11 PRIOR APPLICATION NUMBER: 60/087759  
12 PRIOR FILING DATE: 1998-06-02  
13 PRIOR APPLICATION NUMBER: 60/087827  
14 PRIOR FILING DATE: 1998-06-03  
15 PRIOR APPLICATION NUMBER: 60/088021  
16 PRIOR FILING DATE: 1998-06-04  
17 PRIOR APPLICATION NUMBER: 60/088025  
18 PRIOR FILING DATE: 1998-06-04  
19 PRIOR APPLICATION NUMBER: 60/088026  
20 PRIOR FILING DATE: 1998-06-04  
21 PRIOR APPLICATION NUMBER: 60/088028  
22 PRIOR FILING DATE: 1998-06-04  
23 PRIOR APPLICATION NUMBER: 60/088029  
24 PRIOR FILING DATE: 1998-06-04  
25 PRIOR APPLICATION NUMBER: 60/088030  
26 PRIOR FILING DATE: 1998-06-04  
27 PRIOR APPLICATION NUMBER: 60/088033  
28 PRIOR FILING DATE: 1998-06-04  
29 PRIOR APPLICATION NUMBER: 60/088326  
30 PRIOR FILING DATE: 1998-06-04  
31 PRIOR APPLICATION NUMBER: 60/088167  
32 PRIOR FILING DATE: 1998-06-05  
33 PRIOR APPLICATION NUMBER: 60/088202  
34 PRIOR FILING DATE: 1998-06-05  
35 PRIOR APPLICATION NUMBER: 60/088212  
36 PRIOR FILING DATE: 1998-06-05  
37 PRIOR APPLICATION NUMBER: 60/088217  
38 PRIOR FILING DATE: 1998-06-05  
39 PRIOR APPLICATION NUMBER: 60/088655  
40 PRIOR FILING DATE: 1998-06-09  
41 PRIOR APPLICATION NUMBER: 60/088734  
42 PRIOR FILING DATE: 1998-06-10  
43 PRIOR APPLICATION NUMBER: 60/088738  
44 PRIOR FILING DATE: 1998-06-10  
45 PRIOR APPLICATION NUMBER: 60/088742  
46 PRIOR FILING DATE: 1998-06-10  
47 PRIOR APPLICATION NUMBER: 60/088810  
48 PRIOR FILING DATE: 1998-06-10  
49 PRIOR APPLICATION NUMBER: 60/088824  
50 PRIOR FILING DATE: 1998-06-10  
51 PRIOR APPLICATION NUMBER: 60/088826  
52 PRIOR FILING DATE: 1998-06-10  
53 PRIOR APPLICATION NUMBER: 60/088858  
54 PRIOR FILING DATE: 1998-06-11  
55 PRIOR APPLICATION NUMBER: 60/088861  
56 PRIOR FILING DATE: 1998-06-11  
57 PRIOR APPLICATION NUMBER: 60/088876  
58 PRIOR FILING DATE: 1998-06-11  
59 PRIOR APPLICATION NUMBER: 60/089105  
60 PRIOR FILING DATE: 1998-06-12  
61 PRIOR APPLICATION NUMBER: 60/089440  
62 PRIOR FILING DATE: 1998-06-16  
63 PRIOR APPLICATION NUMBER: 60/089512  
64 PRIOR FILING DATE: 1998-06-16  
65 PRIOR APPLICATION NUMBER: 60/089514  
66 PRIOR FILING DATE: 1998-06-16  
67 PRIOR APPLICATION NUMBER: 60/089532  
68 PRIOR FILING DATE: 1998-06-17  
69 PRIOR APPLICATION NUMBER: 60/089538  
70 PRIOR FILING DATE: 1998-06-17  
71 PRIOR APPLICATION NUMBER: 60/089598  
72 PRIOR FILING DATE: 1998-06-17  
73 PRIOR APPLICATION NUMBER: 60/089599  
74 PRIOR FILING DATE: 1998-06-17  
75 PRIOR APPLICATION NUMBER: 60/089600  
76 PRIOR FILING DATE: 1998-06-17  
77 PRIOR APPLICATION NUMBER: 60/089653  
78 PRIOR FILING DATE: 1998-06-17  
79 PRIOR APPLICATION NUMBER: 60/089801  
80 PRIOR FILING DATE: 1998-06-18  
81 PRIOR APPLICATION NUMBER: 60/089907  
82 PRIOR FILING DATE: 1998-06-18  
83 PRIOR APPLICATION NUMBER: 60/089908  
84 PRIOR FILING DATE: 1998-06-18  
85 PRIOR APPLICATION NUMBER: 60/089947  
86 PRIOR FILING DATE: 1998-06-19  
87 PRIOR APPLICATION NUMBER: 60/089948  
88 PRIOR FILING DATE: 1998-06-19  
89 PRIOR APPLICATION NUMBER: 60/089952  
90 PRIOR FILING DATE: 1998-06-19  
91 PRIOR APPLICATION NUMBER: 60/090246  
92 PRIOR FILING DATE: 1998-06-22  
93 PRIOR APPLICATION NUMBER: 60/090252  
94 PRIOR FILING DATE: 1998-06-22  
95 PRIOR APPLICATION NUMBER: 60/090254  
96 PRIOR FILING DATE: 1998-06-22  
97 PRIOR APPLICATION NUMBER: 60/090349  
98 PRIOR FILING DATE: 1998-06-23  
99 PRIOR APPLICATION NUMBER: 60/090355  
100 PRIOR FILING DATE: 1998-06-23  
101 PRIOR APPLICATION NUMBER: 60/090429  
102 PRIOR FILING DATE: 1998-06-24  
103 PRIOR APPLICATION NUMBER: 60/090431  
104 PRIOR FILING DATE: 1998-06-24  
105 PRIOR APPLICATION NUMBER: 60/090435  
106 PRIOR FILING DATE: 1998-06-24  
107 PRIOR APPLICATION NUMBER: 60/090444  
108 PRIOR FILING DATE: 1998-06-24  
109 PRIOR APPLICATION NUMBER: 60/090445  
110 PRIOR FILING DATE: 1998-06-24  
111 PRIOR APPLICATION NUMBER: 60/090472  
112 PRIOR FILING DATE: 1998-06-24  
113 PRIOR APPLICATION NUMBER: 60/090535  
114 PRIOR FILING DATE: 1998-06-24  
115 PRIOR APPLICATION NUMBER: 60/090540  
116 PRIOR FILING DATE: 1998-06-24  
117 PRIOR APPLICATION NUMBER: 60/090542  
118 PRIOR FILING DATE: 1998-06-24  
119 PRIOR APPLICATION NUMBER: 60/090557  
120 PRIOR FILING DATE: 1998-06-24  
121 PRIOR APPLICATION NUMBER: 60/090676  
122 PRIOR FILING DATE: 1998-06-25  
123 PRIOR APPLICATION NUMBER: 60/090678  
124 PRIOR FILING DATE: 1998-06-25  
125 PRIOR APPLICATION NUMBER: 60/090690  
126 PRIOR FILING DATE: 1998-06-25  
127 PRIOR APPLICATION NUMBER: 60/090694  
128 PRIOR FILING DATE: 1998-06-25  
129 PRIOR APPLICATION NUMBER: 60/090695  
130 PRIOR FILING DATE: 1998-06-25  
131 PRIOR APPLICATION NUMBER: 60/090696  
132 PRIOR FILING DATE: 1998-06-25  
133 PRIOR APPLICATION NUMBER: 60/090862  
134 PRIOR FILING DATE: 1998-06-26  
135 PRIOR APPLICATION NUMBER: 60/090863  
136 PRIOR FILING DATE: 1998-06-26  
137 PRIOR APPLICATION NUMBER: 60/091360  
138 PRIOR FILING DATE: 1998-07-01  
139 PRIOR APPLICATION NUMBER: 60/091478  
140 PRIOR FILING DATE: 1998-07-02  
141 PRIOR APPLICATION NUMBER: 60/091544  
142 PRIOR FILING DATE: 1998-07-01  
143 PRIOR APPLICATION NUMBER: 60/091519  
144 PRIOR FILING DATE: 1998-07-02  
145 PRIOR APPLICATION NUMBER: 60/091626  
146 PRIOR FILING DATE: 1998-07-02

```
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Query Match          100.0%; Score 1608; DB 9; Length 293;
Best Local Similarity 100.0%; Pred. No. 9.1e-150;
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHESNTVPSGNSODLGAGAGEDARS 60
Db 1 MATARPPMMWVLCALITALLIGVTEHVLANNDVSCDHESNTVPSGNSODLGAGAGEDARS 60
QY 61 DSSSRRIINGSDCDMHTOPWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
Db 61 DSSSRRIINGSDCDMHTOPWQAALLRPNQLYCGAVLVHPQWLLTAHCRKKVFRVRLGH 120
QY 121 YSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLIKLNRIRPTKDVPRINVSCHPCS 180
Db 121 YSLSPVYESGQMGFGVKSIPHPGYSHPGHNDMLIKLNRIRPTKDVPRINVSCHPCS 180
QY 181 AGTKCLVSGWGTTSPOVHPKVLQCLNLSVLSQKRCEDAYPRQIDDDTFMFCAGDKAGRDS 240
Db 181 AGTKCLVSGWGTTSPOVHPKVLQCLNLSVLSQKRCEDAYPRQIDDDTFMFCAGDKAGRDS 240
QY 241 CQDSGGPVVNCNGSLQGLVSGWDYPCARPNGPVVTVNLCKFTKWIQETIQANS 293
Db 241 CQDSGGPVVNCNGSLQGLVSGWDYPCARPNGPVVTVNLCKFTKWIQETIQANS 293

RESULT 15
US-09-992-598-309
; Sequence 309, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
```

;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360

;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 100.0%; Score 1608; DB 9; Length 293;  
Best Local Similarity 100.0%; Pred. No. 9.1e-150;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MATARPPMMVLCALITALLGVTEHVLANNDVSCDHPSNTVPSSNQDLGAGAGEDARS 60  
Db 1 MATARPPMMVLCALITALLGVTEHVLANNDVSCDHPSNTVPSSNQDLGAGAGEDARS 60  
  
Qy 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
Db 61 DDSSRIINGSDCDMHTQWQAALLRPNOLYCGAVLVHPQWLLTAACHCRKKVFRVLGH 120  
  
Qy 121 YSLSPVYESQQMFQGVKSIHPHGYSHPGHSDNMLIKLNRRIPTKDVRLPINVSSHCP 180  
Db 121 YSLSPVYESQQMFQGVKSIHPHGYSHPGHSDNMLIKLNRRIPTKDVRLPINVSSHCP 180  
  
Qy 181 AGTKCLVSGWGTGSPQVHPKVLQCLNISVLQKRCEDAYPROIDTMTFCAGDKAGRDS 240  
Db 181 AGTKCLVSGWGTGSPQVHPKVLQCLNISVLQKRCEDAYPROIDTMTFCAGDKAGRDS 240  
  
Qy 241 CQGDSDGGPVVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCKFTKWIQETIOANS 293  
Db 241 CQGDSDGGPVVVCNGSLQGLVSWGDPYPCARPNRPGVYTNLCKFTKWIQETIOANS 293

Search completed: July 4, 2005, 15:19:50  
Job time : 402 secs

Result	No.	Query		DB	ID	Description
		Score	Match Length			
1	627	39.0	260	2	I56559	neuropilin - mouse
2	612.5	38.1	233	2	A53968	serine proteinase
3	586	36.4	246	1	TRRT1	trypsin (EC 3.4.21
4	582.5	36.2	247	2	A27547	trypsin (EC 3.4.21
5	578	35.9	247	1	TRDG	trypsin (EC 3.4.21
6	571.5	35.5	229	1	TRB0TR	trypsin (EC 3.4.21
7	569.5	35.4	248	2	A55066	trypsin (EC 3.4.21
8	563.5	35.0	247	1	A25852	trypsin (EC 3.4.21
9	562.5	35.0	246	1	TRRT2	trypsin (EC 3.4.21
10	561.5	34.9	231	1	TRPQTR	trypsin (EC 3.4.21
11	560	34.8	247	2	S13813	trypsin (EC 3.4.21
12	558.5	34.7	238	2	S31779	trypsin (EC 3.4.21
13	555.5	34.5	246	1	TRDGC	trypsin (EC 3.4.21
14	550	34.2	244	2	A44284	tissue kallikrein
15	544	33.8	246	2	QJ1472	trypsin (EC 3.4.21
16	538	33.5	261	2	A31136	tissue kallikrein
17	537.5	33.4	246	2	B25528	trypsin (EC 3.4.21
18	536	33.3	259	2	B31136	tissue kallikrein
19	535	33.3	246	2	QJ1471	trypsin (EC 3.4.21
20	532.5	33.1	243	2	A35871	trypsin (EC 3.4.21
21	532.5	33.1	248	2	S55067	trypsin (EC 3.4.21
22	530	33.0	261	1	A32297	semenogelase (EC 3
23	529	32.9	261	2	A34079	tissue kallikrein
24	526	32.7	265	1	KQRPp	tissue kallikrein
25	525.5	32.7	247	1	B25852	trypsin (EC 3.4.21
26	525	32.6	261	2	A29586	tissue kallikrein
27	524.5	32.6	304	2	S33496	trypsin (EC 3.4.21
28	523.5	32.6	242	2	S49489	trypsin (EC 3.4.21
29	523.5	32.6	247	2	S05494	trypsin (EC 3.4.21

Matches	115;	Conservative	39;	Mismatches	82;	Indels	8;	Gaps	5;
Qy	50	LGAGAGEDARSDSSRIISGDCDMHTQFWQAALLRPNQLYCGAVLVHPQWLLTAAHC	109						
Db	10	VGRAVAPFLEDDD--KIVGGYTCPEHSVPYQVS--LNSGYHFCGGSINDQWVVSAAHC	64						
Qy	110	RKKVFRVLGHYSLSVPYESGGQFMQGVKSIPHPGYSHPGHSDNMLIKLNRRIPTKOV	169						
Db	65	YKSRIQVRLGEHNIN-VLEGDEQFINAAKTIKHNPYSSTWLTNNDIMLIKLSVPVKLNARV	123						
Qy	170	RPINVSHPESACTKCLVSGWGHTTKSPQVHPFKVLQCLNLSVLSQLCEDAYAPRODDTM	229						
Db	124	APVLPSSACAPAGTCLISGWTNLSGVNPNLLOCDVAPVLSQADCEAAYAGEITSSM	183						

QY	230	FCAGD-KAGRDS	COGDSGGPVCN	SGISGLVSKGDYPCAR	PNP	PGVYVNLCKFTK	WIQET	298
Db	184	ICVGFLEGGKDS	COGDSGGPVCN	QQLQGI	VSWG-YGCAL	PDNPGVYVTKVCN	FVGWIQDT	242
QY	289	IQAN	292					
Db	243	IAAN	246					

RESULT 4

A27547

trypsin (EC 3.4.21.4) precursor, cationic - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004

C;Accession: A27547

R;Fletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C. Biochemistry 26, 3081-3086, 1987

A;Title: Isolation and characterization of a cDNA encoding rat cationic trypsin

A;Reference number: A27547; MUID:87271609; PMID:3607011

A;Residues: 1-247 <FILE>  
 N;Cross-references: UNIPROT:P08426; GB:M16624; NID:G206498; PIDN:AAA41985.1; PI  
 C;Superfamily: trypsin; trypsin homology  
 C;Keywords: calcium binding; hydrolase; protein digestion; serine proteinase  
 F;25-240/domain: trypsin homology <IR>  
 F;31-161,49-65,133-234,140-207,172-186/disulfide bonds: #status predicted  
 F;64,108,201/Active site: His, Asp, Ser #status predicted  
 F;76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted  
 Query Match 36.2%; Score 582.5; DB 2; Length 247;  
 Best Local Similarity 45.6%; Pred. No. 3,8e-44;  
 Query Match 36.2%; Score 582.5; DB 2; Length 247;  
 Best Local Similarity 45.6%; Pred. No. 3,8e-44;

[illegible]

Qy	233	GD-KAGRDS	CQGD	SGG	PV	CN	GS	LQGL	V	SG	DY	P	C	A	R	N	P	G	V	V	T	N	L	C	K	F	T	K	W	I	Q	E	T	I	Q	A	291
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:				
Db	188	GFLEGGK	DS	CQGD	SGG	PV	CN	GS	LQGL	V	SG	DY	P	C	A	R	N	P	G	V	V	T	N	L	C	K	F	T	K	W	I	Q	E	T	I	Q	246
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:				
Qy	292	N	292																																		
Db	247	N	247																																		

RESULT 5

TRDG

trypsin (EC 3.4.21.4) precursor, anionic - dog

N;Alternate names: cationic trypsinogen

C;Species: Canis lupus familiaris (dog)  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004  
C;Accession: A26273  
R;Pinsky, S.D.; LaForge, K.S.; Scheele, G.  
Mol. Cell. Biol. 5, 2669-2676, 1985  
A;Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequence  
A;Reference number: A26273; MUID:86284628; PMID:3841794  
A;Accession: A26273  
A;Molecule type: mRNA  
A;Residues: 1-247 <PIN>  
A;Cross-references: UNIPROT:P06872; GB:M1589; NID:g164094; PIDN:AAA30899.1; PID:g164095  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F;1-15/Domain: signal sequence #status predicted <SIG>  
F;16-23/Domain: activation peptide #status predicted <APT>  
F;24-247/Product: trypsin, anionic #status predicted <TRY>  
F;24-239/Domain: trypsin homology <TRY>  
F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted  
F;63,107,200/Active site: His, Asp, Ser #status predicted  
F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.9%; Score 578; DB 1; Length 247;  
Best Local Similarity 46.1%; Pred. No. 9.5e-44;  
Matches 113; Conservative 40; Mismatches 84; Indels 8; Gaps 5;

Qy 50 LGAGGEDARDDSSRIINGSDCDMHTQWQAALLLRPNOLYCGAVLVHPQWLLTAHC 109  
Db 10 LGAATVATPTDDDD---KIVGGYTCEENSVPYQVS--LNAGYHFCGGLSLSDQWVVAHC 64  
Qy 110 RKVFRVRLGHYLSPVYESQQMFGQVKSIHPGYSHPGHSNDMLIKLRIRPTKDV 169  
Db 65 YKSIQVRLGEYND-VLEGNEQFINSAKVRHPNYSWILDNDMLIKLSSPAVLNARV 123  
Qy 170 RPNVSHCHPSAGTKCLVSGMTTKSPQVHPKVLQCLINISLVSKRCEDAYPRIDDTM 229  
Db 124 ATISLPACAPGPTQCLISGNGTSLSSGTNYPPELLQCLDAPILTAQCEASYPQGITENM 183  
Qy 230 FCAGD-KAGRDSGQSGGPPVVCNGSLQGLVSWGDYPCARNRPGVYTNLCKFKWQET 288  
Db 184 ICAGFLEGGKDCQSGGPPVVCNGELQGLVSWG-YGCAQKNKPGVYTKVNCVFDWIQST 242  
Qy 289 IQANS 293  
Db 243 IAANS 247

RESULT 6  
TRBOTR  
trypsin (EC 3.4.21.4) precursor - bovine  
N;Contains: trypsinogen  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 24-Apr-1984 #sequence\_revision 28-Feb-1986 #text\_change 18-Jul-1997  
C;Accession: A90164; A00946; S08774  
R;Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.  
Biochem. Biophys. Res. Commun. 24, 346-352, 1966  
A;Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.  
A;Reference number: A90164; MUID:67168848; PMID:5967094  
A;Accession: A90164  
A;Molecule type: protein  
A;Residues: 1-57, 'Q', 59-67, 'Q', 69-150, 'N', 152-176, 'N', 178-229 <MIK>  
R;Hartley, B.S.  
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970  
A;Reference number: A93755  
A;Contents: annotation; revisions  
R;Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.  
Biochemistry 14, 1358-1366, 1975  
A;Title: Amino acid sequence of dogfish trypsin.  
A;Reference number: A00950; MUID:75146445; PMID:1092332  
A;Contents: annotation; revisions  
A;Note: the sequence agrees with that shown  
R;Bode, W.; Schwager, P.  
J. Mol. Biol. 98, 693-717, 1975  
A;Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution

A;Reference number: A92954; MUID:76072097; PMID:512  
A;Contents: annotation; x-ray crystallography; binding sites for calcium, substrate, and  
C;Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.  
C;Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a termi  
s pseudotrypsin. A cleavage may also occur after Arg-105.  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F;1-229/Product: trypsinogen #status experimental <ZYM>  
F;1-6/Domain: activation peptide #status experimental <APT>  
F;7-222/Domain: trypsin homology <TRY>  
F;7-131,132-229/Product: alpha-trypsin #status experimental <MPT>  
F;6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental  
F;13-143,31-47,115-216,122-189,154-166,179-203/Disulfide bonds: #status experimental  
F;46,90,183/Active site: His, Asp, Ser #status experimental  
F;58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental  
F;131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match 35.5%; Score 571.5; DB 1; Length 229;  
Best Local Similarity 46.6%; Pred. No. 3.3e-43;  
Matches 108; Conservative 39; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSRIINGSDCDMHTQWQAALLLRPNOLYCGAVLVHPQWLLTAHCRRKVFRLGHY 121  
Db 2 DDDDKIVGGYTCGANTVPYQVS--LNSGYHFCGGLSLNSQWVVAACHYKSGIQVRLGED 59  
Qy 122 SLSPVYESQQMFGQVKSIHPGYSHPGHSNDMLIKLRIRPTKDVPRPNVSHCHPSA 181  
Db 60 NIN-VVSGNEQFTSASKSIHVHPSYNSNTLNDIMLIKLSAASLSNRVASISLPTSCASA 118  
Qy 182 GTKCLVSGMTTKSPQVHPKVLQCLINISLVSKRCEDAYPRQIDDTMFCAGD-KAGRDS 240  
Db 119 GTQCLISGNGTSSGTSPDVLKCLAPILTAQCEASYPQGITSNMFCAGYLEGGKDS 178  
Qy 241 CQSGSGPPVVCNGSLQGLVSWGDYPCARNRPGVYTNLCKFKWQETIQAN 292  
Db 179 CQSGSGPPVVCNGSLQGLVSWGS-GCAQKNKPGVYTKVNCVSVMIKQTIASN 229

RESULT 7  
S55066  
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken  
N;Alternate names: trypsinogen II  
C;Species: Gallus gallus (chicken)  
C;Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C;Accession: S55066; S72347  
R;Wang, K.; Gan, L.; Lee, I.; Hood, L.  
Biochem. J. 307, 471-479, 1995  
A;Title: Isolation and characterization of the chicken trypsinogen gene family.  
A;Reference number: S55065; MUID:95251611; PMID:7733885  
A;Accession: S55066  
A;Molecule type: mRNA  
A;Residues: 1-248 <WAN1>  
A;Cross-references: UNIPROT:Q90629; EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g6039  
A;Experimental source: clone 2-P29  
A;Accession: S72347  
A;Molecule type: DNA  
A;Residues: 1-248 <WAN2>  
A;Cross-references: EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g603907  
A;Experimental source: clone 2-P29  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-25/Domain: activation peptide #status predicted <APT>  
F;26-248/Product: trypsin II #status predicted <MAT>  
F;26-241/Domain: trypsin homology <TRY>  
F;65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 35.4%; Score 569.5; DB 2; Length 248;  
Best Local Similarity 46.1%; Pred. No. 5.4e-43;  
Matches 112; Conservative 38; Mismatches 86; Indels 7; Gaps 5;

Qy 50 LGAGGEDARDDSSRIINGSDCDMHTQWQAALLLRPNOLYCGAVLVHPQWLLTAHC 109  
Db 10 LGAATVATPTDDDD---KIVGGYTCEENSVPYQVS--LNAGYHFCGGLSLSDQWVVAHC 64

```
Db 11 LGAAVAFPGGADD--DKIVGYTTPCHSVYPQVS--LNSGYHFCGSLINSQWYLSAHC 66
Qy 110 RKVFRVRLGHYSLSPVYESGQOMFQGVKSIHPHGYSHPGHNSDLMLIKLNRRIRPTKDV 169
Db 67 YKSIQVRLGYNID-VQEDSEVVRSSVIRHPKYSSITLNDIMLIKLSAIVEYSGADI 125
Qy 170 RPIINVSHCPAGTKCLVSGWGTTKSPQVHPKVLQCLNLSVLQKRCEDAYPRQIDDTM 229
Db 126 QPIALPSSCAKAGTECLISGWNTLSNGYNYPELLQCLNAPILSDQCEQAYPGDITSNM 185
Qy 230 FCAGD-KAGRDSGCGSGPVVCGSLQGLVSWGDYPCARNRPGVYTNLCKFTKWIQET 288
Db 186 ICVGFLEGGKDSGCGSGPVVCGELQGLVSWG-IGCALKYPGVYTKVCNYYVDWIQET 244
Qy 289 IQA 291
Db 245 IAA 247

RESULT 8
A:Species: Homo sapiens (man)
N: trypsin (EC 3.4.21.4) I precursor [validated] - human
C: trypsin (EC 3.4.21.4) I precursor - human
C: Species: Homo sapiens (man)
C: Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C: Accession: A25852; B61066; A43988
R: Emi, M.; Nakamura, Y.; Ogawa, M.; Yamamoto, T.; Nishide, T.; Mori, T.; Matsubara, K.
Gene 41, 305-310, 1986
A: Title: Cloning, characterization and nucleotide sequences of two cDNAs encoding human
A: Reference number: A91544; MUID: 86221712; PMID: 3011602
A: Accession: A25852
A: Molecule type: mRNA
A: Residues: 1-247 <EMI>
A: Cross-references: UNIPROT:P07477; GB:M22612; NID:g521215; PIDN:AAA61231.1; PID:g521216
R: Kilmann, M.; Russick, C.; Marks, W.H.; Borgstroem, A.
Clin. Chim. Acta 184, 31-46, 1989
A: Title: Immunoreactive anionic and cationic trypsin in human serum.
A: Reference number: A61066; MUID: 90091010; PMID: 2598466
A: Accession: B61066
A: Molecule type: protein
A: Residues: 16-43 <KIM>
R: Koivunen, E.; Huhtala, M.L.; Stenman, U.H.
J. Biol. Chem. 264, 14095-14099, 1989
A: Title: Human ovarian tumor-associated trypsin. Its purification and characterization
A: Reference number: A43988; MUID: 89340515; PMID: 2503510
A: Accession: A43988
A: Molecule type: protein
A: Residues: 16-54 <KOI>
A: Experimental source: mucinous ovarian tumor cyst fluid
C: Genetics:
A: Gene: GDB:PRSS1; TRV1
A: Cross-references: GDB:119620; OMIM:276000
A: Map position: 7q35-7q35
A: Note: The human genome contains at least ten trypsin genes or pseudogenes, at least two
C: Superfamily: trypsin; trypsin homology
C: Keywords: hydrolase; pancreas; phosphoprotein; protein digestion; serine proteinase; z
F: 1-15/Domain: signal sequence #status predicted <SIG>
F: 16-246/Product: trypsin I #status experimental <ZYM>
F: 16-246/Product: trypsin I #status predicted <ZYM>
F: 16-246/Product: activation peptide #status experimental <APT>
F: 24-239/Domain: trypsin homology <TRY>
F: 30-160, 48-64, 139-206, 171-185, 196-220/Disulfide bonds: #status predicted
F: 63, 107, 200/Active site: His, Asp, Ser #status predicted
F: 75, 77, 80, 85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.0%; Score 563.5; DB 1; Length 247;
Best Local Similarity 46.4%; Pred. No. 1.8e-42;
Matches 108; Conservative 40; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSSRINGSDCDMHTQWQALLRPNLQYCGAVLVHPQWLLTAHCRKKVFRVLGHY 121
Db 19 DDDDKIVGGYNCSENSVPYQVS--LNSGYHFCGSLINEQWVWSAGHCYKRIQVRLGEH 76
```

```
Qy 122 SLSPVYESGQOMFQGVKSIHPHGYSHPGHNSDLMLIKLNRRIRPTKDVPRPINVSSHCPSA 181
Db 77 NIE-VLEGNQFINAAKIRHPQVDRKTLNDIMLIKLSRAVINARVSTISLTAPPAT 135
Qy 182 GTKCLVSGWGTTKSPQVHPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
Db 136 GTKCLVSGWGTASSGADYDPQLQCLDAPVLSQAKCEASYPGKITSNMFCVGFLEGGKDS 195
Qy 241 CQDGSQSGPVVCGSLQGLVSWGDYPCARNRPGVYTNLCKFTKWIQETIOANS 293
Db 196 CQDGSQSGPVVCGSLQGLVSWGD-GCAQKNRPGVYTKVYVYVVKWIKNTIAANS 247

RESULT 9
TRRT2
trypsin (EC 3.4.21.4) II precursor - rat
N: Alternate names: trypsinogen II
C: Species: Rattus norvegicus (Norway rat)
C: Date: 05-Apr-1983 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C: Accession: A22657; A00949
R: Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A: Title: Structure of two related rat pancreatic trypsin genes.
A: Reference number: A22657; MUID: 85054880; PMID: 6094547
A: Accession: A22657
A: Molecule type: DNA
A: Residues: 1-246 <CRA>
A: Cross-references: UNIPROT:P00763
R: MacDonald, R.J.; Stary, S.J.; Swift, G.H.
J. Biol. Chem. 257, 9724-9732, 1982
A: Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
A: Reference number: A00948; MUID: 82265624; PMID: 6896710
A: Accession: A00949
A: Molecule type: mRNA
A: Residues: 9-246 <MAC>
C: Comment: The trypsin II mRNA is present in much lower quantities than the trypsin I m
C: Genetics:
A: Introns: 14/1; 67/2
A: Superfamily: trypsin; trypsin homology
C: Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F: 1-15/Domain: signal sequence #status predicted <SIG>
F: 16-23/Domain: activation peptide #status predicted <APT>
F: 24-246/Product: trypsin II #status predicted <ENZ>
F: 24-239/Domain: trypsin homology <TRY>
F: 30-160, 48-64, 132-233, 139-206, 171-185/Disulfide bonds: #status predicted
F: 63, 107, 200/Active site: His, Asp, Ser #status predicted
F: 75, 77, 80, 85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 35.0%; Score 562.5; DB 1; Length 246;
Best Local Similarity 45.7%; Pred. No. 2.2e-42;
Matches 106; Conservative 41; Mismatches 80; Indels 5; Gaps 4;

Qy 62 DSSSRINGSDCDMHTQWQALLRPNLQYCGAVLVHPQWLLTAHCRKKVFRVLGHY 121
Db 19 DDDDKIVGGYTCQNSVPYQVS--LNSGYHFCGSLINDQWVWSAAHCYKRIQVRLGEH 76
Qy 122 SLSPVYESGQOMFQGVKSIHPHGYSHPGHNSDLMLIKLNRRIRPTKDVPRPINVSSHCPSA 181
Db 77 NIN-VLSGDEQFINAAKIRHPQVDRKTLNDIMLIKLSRAVINARVATVALPSSCAPA 135
Qy 182 GTKCLVSGWGTTKSPQVHPKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGD-KAGRDS 240
Db 136 GTQCLVSGWNTLSSGVNEPDLQLDAPLQPADCEASYPGKITNDNMVGVGFLEGGKDS 195
Qy 241 CQDGSQSGPVVCGSLQGLVSWGDYPCARNRPGVYTNLCKFTKWIQETIOAN 292
Db 196 CQDGSQSGPVVCGSLQGLVSWG-YGCALPDNPGVYTKVYVYVVKWIKNTIAAN 246

RESULT 10
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N: Contains: trypsinogen
```



C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 24-Apr-1984 #sequence\_revision 24-Apr-1984 #text\_change 09-Jul-2004  
C;Accession: A90641; A90368; A90947  
R;Charles, M.; Koverly, M.; Guidoni, A.; Desnuelle, P.  
Biochim. Biophys. Acta 69, 115-129, 1963  
A;Title: Su le trypsine de porc.  
A;Reference number: A90641  
A;Accession: A90641  
A;Molecule type: Protein  
A;Residues: 1-10 <CHA>  
A;Cross-references: UNIPROT:P00761  
R;Hermanson, M.A.; Erickson, L.H.; Neurath, H.; Walsh, K.A.  
Biochemistry 12, 3146-3153, 1973  
A;Title: Determination of the amino acid sequence of porcine trypsin by sequenator analysis  
A;Reference number: A90368; MUID:73258692; PMID:4738933  
A;Accession: A90368  
A;Molecule type: Protein  
A;Residues: 9-231 <HER>  
A;Note: at position 20, Ile and Val occur alternatively  
A;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym  
F;1-231/Product: trypsinogen #status experimental <ZYM>  
F;1-8/Domain: activation peptide #status experimental <APT>  
F;9-231/Product: trypsin #status experimental <MAT>  
F;9-224/Domain: trypsin homology <TRY>  
F;15-145,33-49,117-218,124-191,156-170,181-205/disulfide bonds: #status predicted  
F;48,92,185/Active site: His, Asp, Ser #status predicted  
F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 34.9%; Score 561.5; DB 1; Length 231;  
Best Local Similarity 45.3%; Pred. No. 2.5e-42;  
Matches 105; Conservative 39; Mismatches 83; Indels 5; Gaps 4;

Qy 62 DSSSRRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAACHKRVFVRLGH 121  
Db 4 DDDDKIVGGYTCANSPYQVS--LNSGSHFCGSLNSQWVSAACHKRSIQVRLGEH 61

Qy 122 SLSVPYSGQMFQGVKSIPIHPGVSHPGHSDMLIKLNRIRPTKDVPRINVSCHPCS 181  
Db 62 NID-VLEGNQFVIAAKIITHPNFNGNTLDNDMLIKLSSPATLNSRVATVSLPRSCAA 120

Qy 182 GTKLVSGWGTTPQVHPFKVQLNLSVLSQKRCEDAYPRQIDDTMFCAAGD-KAGRDS 240  
Db 121 GTECLISGNGTKSGSGSPYSLLOCLKAPVLSDSCKSSYPGQITGNMIVGFLGGKDS 180

Qy 241 CQSDSGGPVVCNGSLQGLVSGDYPFCARPNRPVYTNLCKFTKWIQTIOAN 292  
Db 181 CQSDSGGPVVCNGSLQGLVSGD-YGCAQKNKPGVYTKVCNVVNIQTIAN 231

RESULT 11  
S13813  
trypsin (EC 3.4.21.4) - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 02-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C;Accession: S13813  
R;Le Huerou, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.  
Eur. J. Biochem. 193, 767-773, 1990  
A;Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic  
A;Reference number: S13813; MUID:91065383; PMID:1701147  
A;Accession: S13813  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-247 <HUE>  
A;Cross-references: UNIPROT:Q29463; EMBL:X54703; NID:9829; PIDN:CAA38513.1; PID:9830  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; protein digestion; serine proteinase  
F;24-239/Domain: trypsin homology <TRY>  
F;63,107,200/Active site: His, Asp, Ser #status predicted

Query Match 34.8%; Score 560; DB 2; Length 247;  
Best Local Similarity 47.1%; Pred. No. 3.7e-42;  
Matches 114; Conservative 37; Mismatches 85; Indels 6; Gaps 5;

Qy 53 GAGEDARSDSSRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAACHK 112  
Db 11 GAAVAFPSDD-DDKIVGGYTCANSPYQVS--LNSGSHFCGSLNSQWVSAACHYQY 67

Qy 113 VFRVRLGHYSLSVPYSGQMFQGVKSIPIHPGVSHPGHSDMLIKLNRIRPTKDVPR 172  
Db 68 HIQVRLGEYNID-VLEGGEQPIDASKIIRHPKYSSTWLDNDLILIKLSTPAVINARV 126

Qy 173 NVSSCHPSAGTKLVSGWGTTPQVHPFKVQLNLSVLSQKRCEDAYPRQIDDTMFCA 232  
Db 127 LLPSACASAGTECLISGNGTKSGSPYSLLOCLKAPVLSDSCKSSYPGQITGNMIV 186

Qy 233 GD-KAGRDSGCGSGGPVVCNGSLQGLVSGDYPFCARPNRPVYTNLCKFTKWIQTIO 291  
Db 187 GFLEGGKDSQCGSDSGGPVVCNGSLQGLVSGW-YGCAQKNKPGVYTKVCNVVNIQETIA 245

Qy 292 NS 293  
Db 246 NS 247

RESULT 12  
S31779  
trypsin (EC 3.4.21.4) III precursor - Atlantic salmon (fragment)  
C;Species: Salmo salar (Atlantic salmon)  
C;Date: 22-Nov-1993 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004  
C;Accession: S66657; S31779  
R;Male, R.; Lorens, J.B.; Smalas, A.O.; Torrisen, K.R.  
Eur. J. Biochem. 232, 677-685, 1995  
A;Title: Molecular cloning and characterization of anionic and cationic variants of tryp  
A;Reference number: S66657; MUID:96035908; PMID:7556223  
A;Accession: S66657  
A;Molecule type: mRNA  
A;Residues: 1-238 <MAL>  
A;Cross-references: UNIPROT:P35033; EMBL:X70074; NID:964387; PIDN:CAA49679.1; PID:964388  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase  
F;1-7/Domain: signal sequence (fragment) #status predicted <SIG>  
F;8-15/Domain: activation peptide #status predicted <APT>  
F;16-238/Product: trypsin III #status predicted <MAT>  
F;16-231/Domain: trypsin homology <TRY>  
F;22-152,40-56,124-225,131-198,163-177,188-212/disulfide bonds: #status predicted  
F;55,99,192/Active site: His, Asp, Ser #status predicted

Query Match 34.7%; Score 558.5; DB 2; Length 238;  
Best Local Similarity 44.6%; Pred. No. 4.9e-42;  
Matches 104; Conservative 42; Mismatches 82; Indels 5; Gaps 4;

Qy 61 DSSSRRIINGSCDMDHPTQWQAALLLRPNQLYCGAVLVHPQWLLTAACHKRVFVRLGH 120  
Db 10 DDEDDKIVGGYECRKNSASYQAS--LQSGYHFCGSLNSQWVSAACHKRSIQVRLGE 67

Qy 121 YSLSPYSGQMFQGVKSIPIHPGVSHPGHSDMLIKLNRIRPTKDVPRINVSCHPCS 180  
Db 68 HNIA-VNEGTEQFIDSVKVMHPSYNSRNLNDMLIKLSPASLNSVSTVALPSSCAS 126

Qy 181 AGTKLVSGWGTTPQVHPFKVQLNLSVLSQKRCEDAYPRQIDDTMFCAAGD-KAGRD 239  
Db 127 SGTRCLVSGWNLGSGSSNPDTLRLCLDLPLSSSSCSNAYPGQITSNMFCAGFWEGKD 186

Qy 240 SCQSDSGGPVVCNGSLQGLVSGDYPFCARPNRPVYTNLCKFTKWIQTIOAN 292  
Db 187 SCQSDSGGPVVCNGSLQGLVSGW-YGCAQKNKPGVYTKVCNVYVNIQETIA 238

RESULT 13  
TRDGC  
trypsin (EC 3.4.21.4) precursor, cationic - dog  
N;Alternate names: cationic trypsinogen  
C;Species: Canis lupus familiaris (dog)  
C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004  
C;Accession: S26273



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 4, 2005, 15:07:33 ; Search time 71 seconds  
(without alignments)  
2113.228 Million cell updates/sec

Title: US-09-936-271c-14  
Perfect score: 1608  
Sequence: 1 MATARPPMWWLCAITALL.....VVTNLCFKTKWQETIQANS 293

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	293	1 KLK5 HUMAN	Q9Y337 homo sapien
2	1475	91.7	269	2 Q8IU55	Q8IU55 homo sapien
3	1158.5	72.0	293	2 Q9D140	Q9D140 mus musculus
4	1074	66.8	204	2 Q6S9W8	Q6S9W8 homo sapien
5	683	42.5	276	2 Q8CGR6	Q8CGR6 mus musculus
6	676	42.0	254	2 Q9XSN6	Q9XSN6 sus scrofa
7	671	41.7	254	1 KLK4 HUMAN	Q9Y5K2 homo sapien
8	654	40.7	246	2 Q6P3Z0	Q6P3Z0 mus musculus
9	650.5	40.5	249	2 Q9QYN4	Q9QYN4 mus musculus
10	650.5	40.5	250	1 KLBK HUMAN	Q9UBX7 homo sapien
11	650.5	40.5	250	2 Q632F2	Q632F2 rattus norv
12	650.5	40.5	276	2 Q8QYN3	Q8QYN3 m hipposas
13	643.5	40.0	260	2 Q8TW69	Q8TW69 homo sapien
14	641.5	39.9	256	2 Q6IE12	Q6IE12 rattus norv
15	641.5	39.9	260	1 KLBK HUMAN	Q60259 homo sapien
16	638	39.7	255	2 Q9JIS2	Q9JIS2 mus musculus
17	636.5	39.6	277	1 KLBK HUMAN	Q9UKR3 homo sapien
18	634	39.4	255	2 Q9Z0M1	Q9Z0M1 mus musculus
19	629	39.1	275	2 Q8IXD7	Q8IXD7 homo sapien
20	627.5	39.0	260	1 NRPN RAT	Q88780 rattus norv
21	627	39.0	260	1 NRPN MOUSE	Q61955 mus musculus
22	623.5	38.8	251	2 Q6B089	Q6B089 homo sapien
23	620.5	38.6	251	1 KLBK HUMAN	Q9P0G3 homo sapien
24	612.5	38.1	242	2 Q8OV84	Q8OV84 mus musculus
25	612.5	38.1	250	2 Q8CGR5	Q8CGR5 mus musculus
26	612.5	38.1	253	1 KLBK HUMAN	P49862 homo sapien
27	601.5	37.4	253	2 Q8NSN9	Q8NSN9 homo sapien
28	599.5	37.3	250	1 KLBK HUMAN	Q9UKQ9 homo sapien
29	594	36.9	249	2 Q9IVE3	Q9IVE3 m thymopsin
30	586.5	36.5	247	2 Q9CPN9	Q9CPN9 m mus muscu
31	586	36.4	246	1 TRY1 RAT	P00762 rattus norv

32	585.5	36.4	258	2 Q63ZP4	Q63ZP4 rattus norv
33	582.5	36.2	247	1 TRY3 RAT	P08426 rattus norv
34	580.5	36.1	247	2 Q9D7Y7	Q9D7Y7 mus musculus
35	578.5	36.0	255	2 TRYJG6	P71196 saguinus oe
36	578	35.9	247	1 TRY2 CANFA	P06872 canis famil
37	574.5	35.7	205	2 Q96JE2	Q96JE2 homo sapien
38	574	35.7	234	2 Q9CV76	Q9CV76 mus musculus
39	574	35.7	243	1 TRY1 BOVIN	P00760 bos taurus
40	574	35.7	255	2 Q6GNJ2	Q6GNJ2 xenopus lae
41	573.5	35.7	237	2 Q6GIJ5	Q6GIJ5 struthio ca
42	573	35.6	244	1 TRY2 XENLA	P70059 xenopus lae
43	573	35.6	248	2 Q7SZT1	Q7SZT1 xenopus lae
44	569.5	35.4	248	1 TRY3 CHICK	Q90629 gallus gall
45	569.5	35.4	255	2 Q6ISIO	Q6ISIO homo sapien

ALIGNMENTS

RESULT 1  
ID KLK5\_HUMAN STANDARD; PRT; 293 AA.  
AC Q9Y337; Q9HBG8;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 5 precursor (EC 3.4.21.-) (Stratum corneum tryptic enzyme)  
DE (Kallikrein-like protein 2) (KLK-L2) (UNQ570/PRO1132).  
GN Name=KLK5; Synonyms=SCTE2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Stratum Corneum;  
RX MEDLINE=99445563; PubMed=10514489; DOI=10.1074/jbc.274.42.30033;  
RA Bratteand M., Egelrud T.;  
RT "Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation.";  
RT J. Biol. Chem. 274:30033-30040(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20118156; PubMed=10652563;  
RA Yousef G.M., Luo L.-Y., Diamandis E.P.;  
RT "Identification of novel human kallikrein-like genes on chromosome 19q13.3-q13.4.";  
RT Anticancer Res. 19:2843-2852(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepker B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RT Gene 257:119-130(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J., Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P., Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S., Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J., Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J., Seshagiri S., Simmons I., Singh J., Smith V., Stinson J., Vagts A., Vandlen R., Watanabe C., Wleand R., Woods K., Xie W.-H., Yansura P., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I., Godowski P., Gray A.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";  
RT Genome Res. 13:2265-2270(2003).

[5]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Ovary;  
 RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschuld S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: May be involved in desquamation.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed in skin, breast, brain and testis.  
 CC -!- SIMILARITY: Belongs to the peptidase S1 family, Kallikrein  
 CC subfamily.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AF168768; AAF03101.1; -;  
 DR EMBL; AF135028; AAD26429.1; -;  
 DR EMBL; AF243527; AAG33358.1; -;  
 DR EMBL; AY359010; AAQ89369.1; -;  
 DR EMBL; BC008036; AAH08036.1; -;  
 DR HSSP; P00760; 1EZX.  
 DR MEROPS; S01.017; -;  
 DR Genew; HGNC:6366; KLK5.  
 DR MW; 605643; -;  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
 DR GO; GO:0008544; P:epidermal differentiation; TAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Hydrolyase; Serine protease; Signal.  
 FT SIGNAL 1 22 Potential.  
 FT CHAIN 23 293 Kallikrein 5.  
 FT ACT\_SITE 108 108 Charge relay system (By similarity).  
 FT ACT\_SITE 153 153 Charge relay system (By similarity).  
 FT ACT\_SITE 245 245 Charge relay system (By similarity).  
 FT DISULFID 73 206 By similarity.  
 FT DISULFID 93 109 By similarity.  
 FT DISULFID 178 279 By similarity.  
 FT DISULFID 185 251 By similarity.  
 FT DISULFID 217 231 By similarity.  
 FT DISULFID 241 266 By similarity.  
 FT CARBOHYD 69 69 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 173 173 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 208 208 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 252 252 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 25 56 Missing (in Ref. 3).  
 SQ SEQUENCE 293 AA; 32020 MW; D92C92F5609E5946 CRC64;  
 Query Match 100.0%; Score 1608; DB 1; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-139;  
 Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MATATPPMMVLCALITALLLGVTGHEVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60  
 DB 1 MATATPPMMVLCALITALLLGVTGHEVLANNVSCDHPSTNTPVSGSNQDLGAGAGEDARS 60  
 QY 61 DSSSRRIINGSDCDMHTQPMQALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
 DB 61 DSSSRRIINGSDCDMHTQPMQALLLRPNOLYCGAVLVHPQWLLTAHCRKKVFRVLGH 120  
 QY 121 YSLSPVYESGQMPQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPCS 180  
 DB 121 YSLSPVYESGQMPQGVKSIHPGYSHPGHNSDMLIKLNRRIRPTKDVPRINVSCHPCS 180  
 QY 181 AGTKCLVSGHGTGKSPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 DB 181 AGTKCLVSGHGTGKSPQVHPFKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240  
 QY 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPNNRPGVYTNLCFKTKWIQETIQANS 293  
 DB 241 CQDSDGGPVVCGSLQGLVSWGDYPCARPNNRPGVYTNLCFKTKWIQETIQANS 293  
 RESULT 2  
 Q8IU55 PRELIMINARY; PRT; 269 AA.  
 ID Q8IU55  
 AC Q8IU55, 2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Kallikrein 5 variant (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovarian cancer;  
 RX MEDLINE=22623266; PubMed=12738725;  
 RA Dong Y., Kaushal A., Brattsand M., Nicklin J., Clements J.A.;  
 RT "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer  
 RL produces novel variants with potential as cancer biomarkers.";  
 Clin. Cancer Res. 9:1710-1720(2003).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; AF435980; AAN63606.1; -;  
 DR EMBL; AF435981; AAN63607.1; -;  
 DR HSSP; P00760; 1EZX.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolyase; Protease; Serine protease.  
 FT NON\_TER 269 269  
 SQ SEQUENCE 269 AA; 29226 MW; D3C1E1FBEEA3634C CRC64;  
 Query Match 91.7%; Score 1475; DB 2; Length 269;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-127;

```
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGSNODLGGAGEDARS 60
Db 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGSNODLGGAGEDARS 60
QY 61 DSSSRIRNGSDCDMHTQWQAALLRPNOALYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
Db 61 DSSSRIRNGSDCDMHTQWQAALLRPNOALYCGAVLVHPQWLLTAACHCKKVFVRVLGH 120
QY 121 YSLSPVYESGQOMFGVKSIPHPGYSHPGSHNDMLKLNRRIRPTKQDVRPINVSSHCP 180
Db 121 YSLSPVYESGQOMFGVKSIPHPGYSHPGSHNDMLKLNRRIRPTKQDVRPINVSSHCP 180
QY 191 AGTKCLVSGWGTTKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
Db 191 AGTKCLVSGWGTTKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGRDS 240
QY 241 CQGSQGGPVVCGNSLQGLVSGDYPVPCARP 269
Db 241 CQGSQGGPVVCGNSLQGLVSGDYPVPCARP 269
RESULT 3
Q9D140
ID Q9D140 PRELIMINARY; PRT; 293 AA.
AC Q9D140;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE library, clone:110030019 product:weakly similar to KALLIKREIN 5 (EC
DE 3.4.21.-) (STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN 2)
DE (KIK-L2).
GN Name=1110030019Rik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
```

```
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,
RA Suni N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AK003996; BAB23113.1; -
DR HSSP; P00760; 1BZX.
DR MEROPS; S01.418; -
DR MGB; MGI:1915916; 1110030019Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 293 AA; 31908 MW; ED1F45D8226FE911 CRC64;
Query Match 72.0%; Score 1158.5; DB 2; Length 293;
Best Local Similarity 69.7%; Pred. No. 8.6e-98;
Matches 205; Conservative 41; Mismatches 45; Indels 3; Gaps 2;
QY 1 MATATPPMMVLCALITALLGLVTEHVLANDVSCDHPSNTVPVSGSNODLG--AGAGEDA 58
Db 1 MARTGHPWKWAMATLITTLVLGVSEPVLAGDVSSCDNPSGTEPSTGTRDLSTDSKSGDT 60
QY 59 RSDSSSRIRNGSDCDMHTQWQAALLRPNOALYCGAVLVHPQWLLTAACHCKKVFVRVL 118
Db 61 RS-DSSSRIRNGSDCDMHTQWQAALLRPNOALYCGAVLVHPQWLLTAACHCKKVFVRVL 119
QY 119 GHYSLSPVYESGQOMFGVKSIPHPGYSHPGSHNDMLKLNRRIRPTKQDVRPINVSSHC 178
Db 120 GHHSMSPVYESGQOMFGVKSIPHPGYSHPGSHNDMLKLNRRIRPTKQDVRPINVSSHC 179
QY 179 PSAGTKCLVSGWGTTKSPQVHFPRKVLQCLNLSVLQKRCEDAYPRQIDDTMFCAGDKAGR 238
Db 180 ATEGTRCVSGWGTTSSSHNNFPKVLQCLNITLSEBCKNSYFQIDKTMFCAGDEGR 239
QY 239 DSCQSGGPPVVCNGSLQGLVSGDYPVPCARPVRGVYTNLCKFKTKWIOETQAN 292
Db 240 DSCQSGGPPVVCNGSLQGLVSGDYPVPCARPVRGVYTNLCKFKTKWIOETQAN 293
RESULT 4
Q6S9W8
ID Q6S9W8 PRELIMINARY; PRT; 204 AA.
AC Q6S9W8;
```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Kallikrein 5 isoform 3 preproprotein.
GN Name=KUK5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Michael I.P., Kurlender L., Du D.C., Diamandis E.P.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY461805; AAR23814.1; -.
DR HSP; P36368; IAO5.
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Peptidase_S1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 204 AA; 22100 MW; C9FFDA022246BD52 CRC64;

Query Match 66.8%; Score 1074; DB 2; Length 204;
Best Local Similarity 98.5%; Pred. No. 3.1e-90;
Matches 198; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATATPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNDLGAGEDARS 60
Db 1 MATATPPMMVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNDLGAGEDARS 60

Qy 61 DSSSRIRINGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAACRKKVFRVLGH 120
Db 61 DSSSRIRINGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAACRKKVFRVLGH 120

Qy 121 YSLSPVYESGQMGFGVKSIPHPGSHPGNSNDLMLIKLNRIRPTKDVRLPINSVSHCP 180
Db 121 YSLSPVYESGQMGFGVKSIPHPGSHPGNSNDLMLIKLNRIRPTKDVRLPINSVSHCP 180

Qy 181 AGTKCLVSGWGTTPSQVHFP 201
Db 181 AGTKCLVSGWGTTPSQGEC 201

RESULT 5
ID Q8CGR6 PRELIMINARY; PRT; 276 AA.
AC Q8CGR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Glandular kallikrein KUK13.
GN Name=Kuk13;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Olsson A.Y., Lundwall A.;
RX MEDLINE=2325484; PubMed=12437987; DOI=10.1016/S0006-291X(02)02629-3;
RT "Organization and evolution of the glandular kallikrein locus in Mus musculus.";
RL Biochem. Biophys. Res. Commun. 299:305-311(2002).
RN [2]
RP SEQUENCE FROM N.A.

```

---

```

RA Adams M., Mural R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY152432; AAN7820.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.306; -.
DR MGD; MGI:95292; K1K13.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Peptidase_S1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30360 MW; 33E189C67492FDCA CRC64;

Query Match 42.5%; Score 683; DB 2; Length 276;
Best Local Similarity 46.2%; Pred. No. 3e-54;
Matches 133; Conservative 50; Mismatches 75; Indels 30; Gaps 8;

Qy 9 MWVLCALITALLGVTEHVLANNVSCDHPSTNTVPSGNDLGAGEDARSDDSSRII 68
Db 1 MWPLVATIACLTALSE-----GISRDYK--ILNGTN-----GTSGLP 38

Qy 69 NGSDCMHTOPWQAALLRNQLYCGAVLHPQWLLTAACRKKVFRVLGHYSLSPVYE 128
Db 39 GGYTCLPHSQPWQAALLIR-GRLLCGVLVHPKWLTAACRKGDTVHLGKHALGRV-E 96

Qy 129 SGQMGFGVKSIPHPGV---SHPGNSNDLMLIKLNRIRPTKDVRLPINSVSHCP 183
Db 97 NGEQMEVRSIPHPYQVPTHLNHDHIMLELSPVQLSSHVRLTKLSADCLPTGT 156

Qy 184 KCLVSGWGTTPSQVHFPKVLQCLINISVLQKRCEDAYPRQIDTDFMFCAGDK-AGRDSCQ 242
Db 157 CCRVSGWGTTPSQVNYPKTLQCANIELRSDDECRQVYPGKITANMLCAGTKEGGKDSCE 216

Qy 243 GDGGPVCNGLSVGWDYPCARPNGVYTNLCKFTKWIQETIQ 290
Db 217 GDSGGLICNGKLYGIISWGDFGCGQPNRPGVYTRVSKYLRIIR 264

RESULT 6
ID Q9XSN6 PRELIMINARY; PRT; 254 AA.
AC Q9XSN6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Enamel matrix serine proteinase 1 precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98126310; PubMed=9465170;
RA Simmer J.P., Fukae M., Tanabe T., Yamakoshi Y., Uchida T., Xue J.,
RA Margolis H.C., Shimizu M., Dehart B.C., Hu C.-C., Bartlett J.D.;
RT "Purification, characterization, and cloning of enamel matrix serine proteinase 1.";
RL J. Dent. Res. 77:377-386(1998).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; U76256; AAB94638.1; -.
DR HSP; P00760; IEZX.
DR MEROPS; S01.251; -.
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

```

DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SMO0020; TRYP SEC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease; Signal.  
FT SIGNAL 31 24 Potential.  
FT CHAIN 31 254 enamel matrix serine proteinase 1.  
SQ SEQUENCE 254 AA; 27235 MW; FD40EF85664406F1 CRC64;  
Query Match 42.0%; Score 676; DB 2; Length 254;  
Best Local Similarity 47.4%; Pred. No. 1.2e-53;  
Matches 139; Conservative 35; Mismatches 79; Indels 40; Gaps 6;  
QY 1 MATARPPMMWVLCALITALLGVTEHVLNNDVSCDHPSTNTPVPSGNDQDLAGAGEDARS 60  
DB 1 MTRAGNPGWEL-----GYLLIGIT-----GSLAFINGG----- 29  
QY 61 DSSSRIRSGDCMHTQWQAALLRNQLYCGAVLVHPQWLLTAACRKKVFRVLGH 120  
DB 30 -----HIINGEDCNPHSQPQAALFLE-DDFCGGVLVHPQWLSAAHCFQNSVYIGLGL 83  
QY 121 YSLSPVYVESGQMPGQVKSIHPGYSYHPCGHSNDMLIKLNRIRPTKDYRPIINVSSHCP 180  
DB 84 HNLPEQEPGQSMQMEASLSIHPNEYNEPMSANDMLIKLKEVSLSVTDVTRNISVVSQCP 143  
QY 181 AGTKLVSGMGWTKSPQVHPFKVLQCLNISVLRSOKRCEDAYPRQIDDTMPCA-GDKACRD 239  
DB 144 PGDSCLVSGWRLASGRLL--PQVLCVNISVASEVCKARGPVYHPSPNFCAGGQDQKD 201  
QY 240 SCQDSGSPVYVNCVSLQGLVSGWYPCARPNNRPGVYTNLCRFTKWIQTIQAN 292  
DB 202 SCHGDSGGPLICNGSLQGLVSGWYPCARPNNRPGVYTNLCRFTDWTQTIQAS 254  
RESULT 7  
KLK4 HUMAN STANDARD; PRT; 254 AA.  
ID KLK4 HUMAN O95K2; O95K26; O9UBJ6;  
AC O95K2; O95K26; O9UBJ6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Kallikrein 4 precursor (EC 3.4.21.-) (Protease) (Kallikrein-like protein 1) (KLK-L1) (Enamel matrix serine proteinase 1).  
GN Name=KLK4; Synonyms=EMSP1, PRSS17, PSTS;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=991179024; PubMed=10077646; DOI=10.1073/pnas.96.6.3114;  
RA Nelson P.S., Gan L., Ferguson C., Moss P., Gellinas R., Hood L., Wang K.;  
RT "Molecular cloning and characterization of prostate, an androgen-regulated serine protease with prostate-restricted expression.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:3114-3119 (1999).  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99413477; PubMed=10485467;  
RA Yousef G.M., Obiezu C.V., Luo L.-Y., Black M.H., Diamandis E.P.;  
RT "Protease/KLK-L1 is a new member of the human kallikrein gene family, is expressed in prostate and breast tissues, and is hormonally regulated.";  
RL Cancer Res. 59:4252-4256 (1999).  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99367447; PubMed=10438493; DOI=10.1074/jbc.274.33.23210;  
RA Stephenson S.A., Verity K., Ashworth L.K., Clements J.A.;  
RT "Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3-13.4.";  
RL J. Biol. Chem. 274:23210-23214 (1999).  
RP SEQUENCE FROM N.A.  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J., Moss P., Paepers B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130 (2000).  
RP SEQUENCE FROM N.A. AND CHARACTERIZATION  
RX MEDLINE=20323211; PubMed=10863090; DOI=10.1016/S0378-1119(00)00203-1;  
RA Hu J.C.-C., Zhang C., Sun X., Yang Y., Cao X., Ryu O., Simmer J.P.;  
RT "Characterization of the mouse and human PRSS17 genes, their relationship to other serine proteases, and the expression of PRSS17 in developing mouse incisors.";  
RL Gene 251:1-8 (2000).  
RP SEQUENCE FROM N.A.  
RX PubMed=15057824; DOI=10.1038/nature02399;  
RA Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J., Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M., Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E., Caenepeel S., Carrano A.V., Caoille C., Chan Y.M., Christensen M., Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Dettler J.C., Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M., Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T., Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H., Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S., Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J., Nelson K., Nolan M., Ovcharenko I., Pittluck S., Pollard M., Popkie A.P., Predki P., Quan Q., Ramirez L., Rash S., Retterer J., Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D., Szekak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A., Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I., Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E., Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M., Rubin E.M., Lucas S.M.;  
RT "The DNA sequence and biology of human chromosome 19.";  
RL Nature 428:529-535 (2004).  
RP SEQUENCE OF 22-254 FROM N.A.  
RX Simmer J.P., Ryu O.H., Qian Q., Zhang C., Cao X., Sun X., Hu C.-C.;  
RT "Cloning and characterization of a cDNA encoding human EMSP1.";  
RL (In) Goldberg M. (eds.);  
RP Chemistry and biology of mineralized tissues, pp.1-1, American Academy of Orthopaedic Surgeons, Vitell (2000).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Expressed in prostate.  
CC -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein subfamily.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF113140; AAD21580.1; -  
DR EMBL; AF113141; AAD21581.1; -  
DR EMBL; AF135023; AAD26424.2; -  
DR EMBL; AF148532; AAD38019.1; -  
DR EMBL; AF243527; AAG33357.1; -









```

[7]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Testis;
MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altshuler S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T.I., Wang J., Haiech F.,
Rachakonda L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
-!- FUNCTION: Possible multifunctional protease. Efficiently cleaves
bz-Phe-Arg-4-methylcoumaryl-7-amide, a kallikrein substrate, and
weakly cleaves other substrates for kallikrein and trypsin.
-!- SUBCELLULAR LOCATION: Secreted.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q9UBX7-1; Sequence=Displayed;
Name=2;
IsoId=Q9UBX7-2; Sequence=VSP 005402;
-!- TISSUE SPECIFICITY: Expressed in brain, skin and prostate. Isoform
1 is expressed preferentially in brain; isoform 2 in prostate.
-!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AB012917; BAA33404.1; ALT_INIT.
EMBL; AB013730; BAA88713.1; -.
EMBL; AB041036; BAA96797.1; -.
EMBL; AF164623; AAD47815.1; -.
EMBL; AF243527; AAG33364.1; -.
EMBL; AY359014; AAG89373.1; -.
EMBL; AC011473; AAG23257.1; -.
EMBL; BC022068; AAH22068.1; -.
HSP; P00760.1; IEZX.
MEROPS; S01.257; -.
Genew; HGNC:6359; KUK11.
H-invDB; HIX0015375; -.
MIM; 604434; -.
GO; GO:0008236; F:serine-type peptidase activity; TAS.
InterPro; IPR009003; Pept_Ser_Cys.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00089; Trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Alternative splicing; Glycoprotein; Hydrolase; Serine protease;
Signal; Zymogen.
SIGNAL 1 18 Potential.
PROPEP 19 21 Activation peptide (Potential).

```

```

FT CHAIN 22 250 Kallikrein 11.
FT ACT SITE 62 62 Charge relay system (By similarity).
FT ACT SITE 110 110 Charge relay system (By similarity).
FT ACT SITE 203 203 Charge relay system (By similarity).
FT DISULFID 28 163 By similarity.
FT DISULFID 47 63 By similarity.
FT DISULFID 135 237 By similarity.
FT DISULFID 142 209 By similarity.
FT DISULFID 174 188 By similarity.
FT DISULFID 199 224 By similarity.
FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 165 165 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 181 181 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 210 210 N-linked (GlcNAc...) (Potential).
FT VARSPPLIC 1 1 M -> MORLRWRDKSSGRGLTAAKPGARSSPLQAM
(in isoform 2).
FT FT /FTId=VSP_005402
SQ SEQUENCE 250 AA; 27466 MW; 192D910BECDC7A56 CRC64;

Query Match 40.5%; Score 650.5; DB 1; Length 250;
Best Local Similarity 50.6%; Pred. No. 2.5e-51;
Matches 118; Conservative 41; Mismatches 67; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDMHTQFWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGHYSLS 124
Db 20 TRIIKGFCEKPHSQWQAALFEKTRLLCGATLAPRWLLTAARCLKPRYIVHLGQHNLQ 78
Qy 125 PVYESGOQMFQGVKSIHPGVS---HPGHSNDLMLKLNRRIRPTKDVRLPINVSSHCPSS 180
Db 79 K-EGCEQTATSEFPFGFNLSLNKDRHNDLMLVKNASPVSIITWAVRLTUSRCVT 137
Qy 181 AGTKCLVSGWGTTKSPQVHPFKVLQCLNI SVLSOKRCEDAYPRQIDDTMFAG-DKAGRD 239
Db 138 AGTSLISGWSSTSPQLRLPHTLRCAITIIIEHKCENAYPGNITDTWVCASVQEGSKD 197
Qy 240 SCQSDSGPVVNGSLQGLVSGWGYPCARNRPQVYTNLCKFTKWIQETIOAN 292
Db 198 SCQSDSGPLVNCQSLQGLIISWGQDPCAITRKPGVYTKVCKYVDWIOBTMKN 250

RESULT 11
Q63ZF2 Q63ZF2 PRELIMINARY; PRT; 250 AA.
AC Q63ZF2; 2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Glandular kallikrein 11 precursor.
GN Name=Klk1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15203212; DOI=10.1016/j.ygeno.2004.01.009;
RA Olsson A.Y., Lilja H., Lundwall A.;
RT "Taxon-specific evolution of glandular kallikrein genes and
RT identification of a progenitor of prostate-specific antigen.";
RL Genomics 84:147-156 (2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001364; DAA05600.1; -.
KW Signal.
FT SIGNAL 1 18 Potential.
FT CHAIN 22 250 glandular kallikrein 11.
SQ SEQUENCE 250 AA; 27630 MW; 247B29D3F0DEF8F1 CRC64;

Query Match 40.5%; Score 650.5; DB 2; Length 250;
Best Local Similarity 50.6%; Pred. No. 2.5e-51;
Matches 118; Conservative 40; Mismatches 68; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDMHTQFWQAALLRPNQLYCGAVLVHPQWLLTAACRKKVFRVLGHYSLS 124

```

```
Db 20 TRIIKGYECRPHSQPQVA-LFQKTRLLCGATLIAPKWLITAAHCKRKHVYLLGEHNL 78
Qy 125 PYESGQOMFGQVKSIPHPGYS----HPGSHNDMLIKLNRRIRPTKDVPRINVSCHPS 180
Db 79 KT-DGCEQRRMATESFPFHPGNSLPNKDHRNDIMLVKSSPAFTTAVRPLTSSLCVT 137
Qy 181 AGTKCLVSGWGTTSQPQVHFPKVLQCLNLSVLSQKRCEDAYPRQIDDDTMFCAG-DKAGRD 239
Db 138 AGTSLISGWGTTSPQRLPHSLRCANVSIIHGKCEKRAIPGNITDMLCASVRKEGKD 197
Qy 240 SCQDGSQGVVNGSLQGLVSGDYPFCARPNRPVYTNLCKRFTKWIQTIQAN 292
Db 198 SCQDGSQGLVNGSLQGLISWGQPCAVTRKPGVYTKVCKYFDWIHEVMRN 250

RESULT 12
Q9QYN3 PRELIMINARY; PRT; 276 AA.
ID Q9QYN3
AC Q9QYN3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hipostatinsin prostate type (Mus musculus adult male tongue cDNA, RIKEN
DE full-length enriched library, clone:2310015108 product:protease,
DE serine, 20, full insert sequence) (Mus musculus adult male tongue
DE cDNA, RIKEN full-length enriched library, clone:2310040P07
DE product:protease, serine, 20, full insert sequence).
GN Name=2310015108Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=20525460; PubMed=11072088; DOI=10.1016/S0167-4781(00)00206-2;
RA Mitsui S., Okui A., Koninami K., Uemura H., Yamaguchi N.;
RT "cDNA cloning and tissue-specific splicing variants of mouse
RT hipostatinsin/TLSP (PRSS20).";
RL Biochim. Biophys. Acta 1494:206-210(2000).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=92729253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
```

```
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=20530913; PubMed=11078861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Suni N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[8]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AB016227; BAB36955.1; -
DR EMBL; AK009360; BAB26241.2; -
DR EMBL; AK009720; BAB26461.2; -
DR HSRF; P00760; IEZX.
DR MEROPS; S01.257; -.
DR MG; MG1:1929977; 2310015108Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 276 AA; 30753 MW; 90BDC03A8AB178D6 CRC64;

Query Match 40.5%; Score 650.5; DB 2; Length 276;
Best Local Similarity 49.8%; Pred. No. 2.8e-51;
Matches 116; Conservative 42; Mismatches 68; Indels 7; Gaps 4;

Qy 65 SRIINGSDCDMHTQPOWQAALLRPOLYCGAVLVHPQWLLTAAHCKRKHVYLLGEHNL 124
Db 46 TRIIKGYECRPHSQPQVA-LFQKTRLLCGATLIAPKWLITAAHCKRKHVYLLGEHNL 104
Qy 125 PYESGQOMFGQVKSIPHPGYS----HPGSHNDMLIKLNRRIRPTKDVPRINVSCHPS 180
Db 105 KT-DGCEQRRMATESFPFHPGNSLPNKDHRNDIMLVKSSPAFTTAVRPLTSSLCVT 163
Qy 181 AGTKCLVSGWGTTSQPQVHFPKVLQCLNLSVLSQKRCEDAYPRQIDDDTMFCAG-DKAGRD 239
Db 164 AGTSLISGWGTTSPQRLPHSLRCANVSIIHGKCEKRAIPGNITDMLCASVRKEGKD 223
Qy 240 SCQDGSQGVVNGSLQGLVSGDYPFCARPNRPVYTNLCKRFTKWIQTIQAN 292
Db 224 SCQDGSQGLVNGSLQGLISWGQPCAVTRKPGVYTKVCKYFDWIHEVMRN 276
```

```
RESULT 13
Q8IW69
ID Q8IW69 PRELIMINARY; PRT; 260 AA.
AC Q8IW69;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Kallikrein 8, isoform 1 preproprotein.
GN Name=KLK8;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.H.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schain J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; BC040887; AAH40887.1; -.
DR HSSP; P00760; IEZ.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 260 AA; 28090 MW; EF5934EB96295660 CRC64;

Query Match 40.0%; Score 643.5; DB 2; Length 260;
Best Local Similarity 42.6%; Pred. No. 1.2e-50;
Matches 126; Conservative 43; Mismatches 80; Indels 47; Gaps 7;

QY 1 MATARP-----PMMVLCALITALLGVTEHVLNNDVSCDHPSNTVPSSNODLGAGAGE 56
DB 1 MCRPRAKWTMFL-----LLLG-----GAWAGH 25

QY 57 DARSDDSSRIINGSDCMHTQFWQAALLLRPNQYCGAVLVHPQWLLTAACRKKVFRV 116
DB 26 SRAQED---KVLGHECQPHSQFWQA--LFQCGQLCGGVLVGGNWLTAACRKKPKYV 81

us-09-936-271c-14.rup
117 RLGHYSLSPYVESGQQMFQGVKSIPIHPGYGH---PGHSNDMLMLKLNRRIRPTKDVVRPN 173
82 RLGDHSLQN-KDGPQEIPVQSIPIHPFCYNSDVDEHNDLMLQLRDQASLGSKVKPIIS 140
174 VSHCFPSAGTKCLVSGMTTKYSQVHPKVLQCLNISVLQKRCEDAYPRQIDTMCAG 233
141 LADHTQPGQKCTVSGMTVTSPRENFPDITLNCAEVAKIFFQKKCEDAYPGQITDVMVCAG 200
234 DKAGRDSGCGSGPVVCGSLVSGDYPCCARPNNRPGVYTNLCFKFTKWIQETI 289
201 SSKGADTCQGDGSGPLVCDGALQITSGSDPCGRSKDPGVYTNICRYLDWIKKII 256

RESULT 14
Q6IE12
ID Q6IE12 PRELIMINARY; PRT; 256 AA.
AC Q6IE12;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Kallikrein 4 precursor.
GN Name=klk4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX PubMed=15060002; DOI=10.1101/gr.1946304;
RA Puente X.S., Lopez-Otin C.;
RT "A genomic analysis of rat proteases and protease inhibitors.";
RL Genome Res. 14:609-622(2004).
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BN000381; CAB51907.1; -.
DR HSSP; P00761; IAKS.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 30 Potential.
FT CHAIN 31 256 kallikrein 4.
SQ SEQUENCE 256 AA; 27925 MW; 8860B91535DB3285 CRC64;

Query Match 39.9%; Score 641.5; DB 2; Length 256;
Best Local Similarity 46.4%; Pred. No. 1.7e-50;
Matches 135; Conservative 29; Mismatches 88; Indels 39; Gaps 5;

QY 1 MATARPMMVLCALITALLGVTEHVLNNDVSCDHPSNTVPSSNODLGAGAGEDARS 60
DB 2 MVTARTPWGFLGYLI-----LEVT-----GSSASS 27

QY 61 DSSSRRIINGSDCMHTQFWQAALLLRPNQYCGAVLVHPQWLLTAACRKKVFRVLGH 120
DB 28 --ISRIIQGQDCLPHSQFWQAALFSEDNAFFCSGLVHPQWLVSAAHCIODSYTVGLGL 85

QY 121 YSLSPVSVESGQQMGVKSIPHPGYGHGPHGNDMLMLKLNRRIRPTKDVVRPNVSSHCP 180
DB 86 HNLEGSQPGGRMLEAHLSIOHPNYNDPSFANDLMLKLNESVMESNTRIRIPVASQCP 145
```

OY 181 AGTKCLVSGWGTTSKSPVHFVKVLQCLNLSVLSOKRCEDAYPRQIDDTMFAGDKAGR-D 239  
DB 146 PGDTCLVSGWGLRNGKL--PSLQCVNLSVASSETCRLLDVYVHLSNFCAGGGPRKD 203  
OY 240 SCQSDGSPVVCVNGSLQSLVSGWGDVPCARPNRPGVYTNLCKFTKWIQTIIQ 290  
DB 204 TCNGDSGGPIVCNRSQSLVSGWGCQCGPGIPSVYTNLCKFTNWIQTIIQ 254

RESULT 15  
ID KX8 HUMAN STANDARD; PRT; 260 AA.  
AC O60259; Q9HCB3; Q9U1L9; Q9UQ47;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Neuropilin precursor (EC 3.4.21.-) (NP) (kallikrein 8) (Ovasin) (Serrine  
DE protease TADG-14) (tumor-associated differentially expressed gene-14  
DE protein) (UNQ283/PRO322).  
GN Name=KLK8; Synonyms=NRPN, PRSS19, TADG14;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Hippocampus;  
RX MEDLINE=98372070; PubMed=9714609; DOI=10.1016/S0378-1119(98)00232-7;  
RA Yoshida S., Taniguchi M., Hirata A., Shiosaka S.;  
RT "Sequence analysis and expression of human neuropilin cDNA and gene.";  
RL Gene 213:9-16(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Brain;  
RX MEDLINE=99203457; PubMed=10102990;  
RA Mitsui S., Tsuruoka N., Yamashiro K., Nakazato H., Yamaguchi N.;  
RT "A novel form of human neuropilin, a brain-related serine protease, is  
RT generated by alternative splicing and is expressed preferentially in  
RT human adult brain.";  
RL Eur. J. Biochem. 260:627-634(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Ovary;  
RX MEDLINE=99413504; PubMed=10485494;  
RA Underwood L.J., Tanimoto H., Wang Y., Shigemasa K., Parmley T.H.,  
RA O'Brien T.J.;  
RT "Cloning of tumor-associated differentially expressed gene-14, a novel  
RT serine protease overexpressed by ovarian carcinoma.";  
RL Cancer Res. 59:4435-4439(1999).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Gan L., Gelinas R., Gown A.M., Moss P., Smith R., Wang K.;  
RT "Molecular cloning and characterization of a novel serine protease,  
RT ovasin, a potential molecular marker for ovarian carcinomas.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;  
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,  
RA Moss P., Paepel B., Wang K.;  
RT "Sequencing and expression analysis of the serine protease gene  
RT cluster located in chromosome 19q13 region.";  
RL Gene 257:119-130(2000).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
RA Vandlen R., Watanabe C., Wleand D., Woods K., Xie M.-H., Yansura D.,

Yi S., Yu P., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
Godowski P., Gray A.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
RT effort to identify novel human secreted and transmembrane proteins: a  
RT bioinformatics assessment.";  
RL Genome Res. 13:2265-2270(2003).  
RN [7]  
RP SEQUENCE OF 1-164 FROM N.A. (ISOFORM 1).  
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Ganes J.,  
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Andreise T., Frankheim M., Attix C., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RT "Sequence analysis of chromosome 19q13.4.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
CC 1- FUNCTION: Suggested to be involved in kindling epileptogenesis and  
CC hippocampal plasticity.  
CC 1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.  
CC 1- SUBCELLULAR LOCATION: Secreted.  
CC 1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=O60259-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=O60259-2; Sequence=VSP\_005401;  
CC 1- TISSUE SPECIFICITY: Isoform 1 is predominantly expressed in the  
CC pancreas while isoform 2 is expressed in adult brain and  
CC hippocampus. Both forms are also found in fetal brain and  
CC placenta. Not detected in kidney, spleen, liver and lung.  
CC 1- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein  
CC subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AB009849; BAA28673.1; -  
CC EMBL; AB012761; BAA28676.1; -  
CC EMBL; AB010780; BAA88684.1; -  
CC EMBL; AB008390; BAA82665.1; -  
CC EMBL; AB008927; BAA82666.1; -  
CC EMBL; AF055982; AAD56050.1; -  
CC EMBL; AF095742; AAD25979.1; -  
CC EMBL; AF095743; AAD29574.1; -  
CC EMBL; AF243527; AAG33361.1; -  
CC EMBL; AY359036; AAQ89395.1; -  
CC EMBL; AC011473; AAG23254.1; -  
CC HSPSP; P00760.1; IEZX.  
CC MEROPS; S01.244; -  
CC Gene; HGNC:6369; KLK8.  
CC MIW; G05644; -  
CC GO; GO:0008236; P:serine-type peptidase activity; TAS.  
CC GO; GO:0007339; P:neurogenesis; TAS.  
CC InterPro; IPR009003; Pept\_Ser\_Cys.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; Trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Alternative splicing; Glycoprotein; Hydrolase; Serine protease;  
KW Signal; Zymogen.  
FT SIGNAL 1 28 Potential.  
FT PROPEP 29 32 By similarity.  
FT CHAIN 33 260 Neuropilin.

